

GenCore version 5.1.4.F5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 16, 2003, 08:59:57 ; Search time 26 Seconds  
(without alignments)  
520.049 Million cell updates/sec

Title: US-10-067-989-1  
Perfect score: 1654  
Sequence: 1 MASLRLFTSTHNSQLLPSSL..... KAVWREPRKRGFFSF95 326

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	799.5	48.3	286	1 MIND_MESVI	Q9MUN5 mesostigma
2	792	47.9	282	1 MIND_CHLVU	P63346 chlorella v
3	694.5	42.0	269	1 MIND_GUITH	O78436 guillardia
4	643	38.9	266	1 MIND_SYNY3	O55900 synochocyst
5	523	31.6	268	1 MIND_BACSU	O01464 bacillus su
6	506.5	30.6	268	1 MIND_HELPY	O25098 helicobacte
7	500.5	30.3	268	1 MIND_HELPY	O25098 helicobacte
8	479	29.0	269	1 MIND_HELPY	O25098 helicobacte
9	475.5	28.7	271	1 MIND_HELPY	O25098 helicobacte
10	460	27.8	269	1 MIND_HELPY	O25098 helicobacte
11	300	18.1	264	1 MIND_HELPY	O25098 helicobacte
12	242	14.6	263	1 MIND_HELPY	O25098 helicobacte
13	214	12.9	263	1 MIND_HELPY	O25098 helicobacte
14	213	12.9	353	1 MIND_HELPY	O25098 helicobacte
15	211	12.8	364	1 MIND_HELPY	O25098 helicobacte
16	209	12.6	368	1 MIND_HELPY	O25098 helicobacte
17	206	12.5	368	1 MIND_HELPY	O25098 helicobacte
18	197.5	11.9	350	1 MIND_HELPY	O25098 helicobacte
19	196.5	11.9	295	1 MIND_HELPY	O25098 helicobacte
20	193.5	11.7	295	1 MIND_HELPY	O25098 helicobacte
21	191	11.5	290	1 MIND_HELPY	O25098 helicobacte
22	190	11.5	381	1 MIND_HELPY	O25098 helicobacte
23	183.5	11.1	253	1 MIND_HELPY	O25098 helicobacte
24	181	10.9	253	1 MIND_HELPY	O25098 helicobacte
25	181	10.9	287	1 MIND_HELPY	O25098 helicobacte
26	177.5	10.7	383	1 MIND_HELPY	O25098 helicobacte
27	173.5	10.5	271	1 MIND_HELPY	O25098 helicobacte
28	173.5	10.5	369	1 MIND_HELPY	O25098 helicobacte
29	165	10.0	257	1 MIND_HELPY	O25098 helicobacte
30	164.5	9.9	267	1 MIND_HELPY	O25098 helicobacte
31	164.5	9.9	267	1 MIND_HELPY	O25098 helicobacte
32	161	9.7	275	1 MIND_HELPY	O25098 helicobacte
33	160.5	9.7	287	1 MIND_HELPY	O25098 helicobacte

34	153.5	9.3	290	1 BCL_RHURU	Q91815 rhodospirill
35	152	9.2	352	1 MRP_BACSU	P50863 bacillus su
36	150.5	9.1	253	1 SOJ_BACHD	O9K5N0 bacillus su
37	149.5	9.0	263	1 YG1_PSEPU	P31856 pseudomonas
38	146.5	8.9	276	1 BCL_CHLIT	O9F714 chlorobium
39	146.5	8.9	293	1 YIA3_YEAST	P40558 saccharomyc
40	144.5	8.7	370	1 MRP_HELPY	P45135 haemophilus
41	143	8.6	320	1 NUB1_HUMAN	P53384 homo sapien
42	140.5	8.5	296	1 NUB1_ALCPA	O44044 alcaligenes
43	140.5	8.5	320	1 NUB1_MOUSE	O9F060 mus musculi
44	140	8.5	253	1 SOJ_TREPA	O83296 treponema p
45	137	8.3	287	1 NIFH_FRAAL	P08925 frankia aln

## ALIGNMENTS

RESULT 1	ID	MIND_MESVI	STANDARD;	PRT;	286 AA.
AC	Q9MUN5;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Putative septum site-determining protein mind.				
GN	MIND.				
OS	Mesostigma viride.				
OG	Chloroplast.				
OC	Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;				
OC	Mesostigmatatales; Mesostigmatataceae; Mesostigma.				
OX	NCBI_TaxID=41882;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=NIES-296;				
RC	MEDLINE=20150907; PubMed=10688199;				
RA	Lemieux C., Otis C., Turmel M.;				
RT	"Ancestral chloroplast genome in Mesostigma viride reveals an early				
RL	branch of green plant evolution.";				
RT	Nature 403:649-652(2000).				
CC	- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE				
CC	- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.				
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation.				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
DR	EMBL; AF161114; AAFA43873.1;				
DR	InterPro; IPR00707; ATPase_Para.				
DR	Pfam; PF00991; Para; 1.				
FT	Cell division; Septation; ATP-binding; Chloroplast.				
KW	NP BIND 25				
SC	SEQUENCE 286 AA; 31627 MW; 250888F822583F9 CXC64;				
Query Match	48.3%; Score 799.5; DB 1; Length 286;				
Best Local Similarity	60.4%; Pred. No. 8.4e-57;				
Matches 166; Conservative 39; Mismatches 61; Indels 9; Gaps 4;					
QY	46 QPFRKELAGET-PRIVITSGKGVGKTTTANGLSARGFVVAIDDLGRINDL 104				
DB	4 QIKKDEKKNSTDTKITVITSGKGVGKTTTANGLSARGFVVAIDDLGRINDL 63				
QY	105 LIGLENRVVTCVENVINSDQDAIVPRKWSNFELICISPKPSKLPMGFGKALEMLV 164				
DB	64 LIGLENRVVTCVENVINSDQDAIVPRKWSNFELICISPKPSKLPMGFGKALEMLV 121				
QY	165 DAKTTPSPSPFTITDPAAGIDAGFTTATATAMAVLVTTTETATLADADRVYGLFCD 224				
DB	122 DSIIRL---NYNFILIDCPAGIDVGFVNAVAPAEAVVVTTPETISIRDAADRVGLTAS 178				

QY 225 GTPDITM:VMPVPTM:VQEDMSVLDVQEM:GILGIVIDESEVISTNPGPLVANK 284  
 DB 179 GIVEVLLVNPVPMI:GNMML:SPVDQEM:GILGALPEDIIVIASTRGQPLVANK 238  
 QY 295 PPLIAGLAEQAMPLVEEDSMKAVMVEEPKRG 319  
 DB 239 KLLTSGISFENMAPRLV--GRKEYLVNLETGNGK 270

## RESULT 2

ID MIND\_CHUVU STANDARD; PRT; 282 AA.  
 AC P56346;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative septum site-determining protein mind.  
 GN MIND.  
 OS Chlorella vulgaris.  
 OC Chloroplast.  
 CC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales,  
 OC Chlorellaceae; Chlorella.  
 OX NCBI\_TaxID=3077;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 PC STRAIN=JAM C-27 / T-173;  
 RX MEDLINE=97303241; PubMed=9159184;  
 RA Waksugi T., Nagai T., Kapoor M., Sugita M., Ito S.,  
 RA Tsuchiura T., Nakashima K., Tsuchiura T., Suzuki Y., Hamada A., Ohta T.,  
 RA Inamura A., Yoshinaga K., Sugitara M.;  
 RT "Complete nucleotide sequence of the chloroplast genome from the  
 RT green alga Chlorella vulgaris: the existence of genes possibly  
 RT involved in chloroplast division."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997)  
 CC FUNCTION: ATPase REQUIRED FOR THE CORRECT PLACEMENT OF THE  
 CC DIVISION SITE (BY SIMILARITY).  
 CC - SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC FMBU, AB001684; BAA57951.1;  
 DR InterPro: IPR000707; ATPase\_Para  
 DR Pfam: PF00991; Para; 1.  
 KW Cell division; Septation; ATP-binding; Chloroplast  
 FT NP BIND 24 31  
 FT SEQUENCE 282 AA; 31013 MW; 94010DD45ABE67 CRC64;

Query Match 47.9%; Score 792, DB 1, Length 282,  
 Best Local Similarity 60.4%; Pred. No. 3 3e-56;  
 Matches 165; Conservative 39; Mismatches 55; Indels 14; Gaps 5;

QY 59 RIVVITSGGKGGKTTTANVGLSLARYGSVAIDADLGRNDLLGLLENRVNTTVE 118  
 DB 17 RVVITSGGKGGKTTTANVGLSLARYGSVAIDADLGRNDLLGLLENRVNTTVE 76  
 QY 119 VINGDCLDQALVDPKPMKSNFELLCTSKPPSKLPMWFGGKALEMVLNLTPTPEGSPDFI 178  
 DB 77 IVGGGCLDQALVDPKPMKSNFELLCTSKPPSKLPMWFGGKALEMVLNLTPTPEGSPDFI 133  
 QY 179 IIDCPAGIDAGFTTATPANEAULTPTDITLRDADPVTSLEGTGIPDIKMTVNPVPT 238  
 DB 132 IIDCPAGIDAGFTTATPANEAULTPTDITLRDADPVTSLEGTGIPDIKMTVNPVPT 191  
 QY 239 DMKIGEDMMSVLDVQEMGLSLGVPEDESEVISTNPGPLVANKPPLAGLAEQAM 298  
 DB 102 DMKIGEDMMSVLDVQEMGLSLGVPEDESEVISTNPGPLVANKPPLAGLAEQAM 251

QY 239 PLV-EQDSMAVWVEEPKRGPF...SFGG 326  
 DB 252 PLV-EQDSMAVWVEEPKRGPF...SFGG 280

## RESULT 3

ID MIND\_GUITH STANDARD; PRT; 269 AA.  
 AC 078436;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative septum site-determining protein mind.  
 GN MIND.  
 OS Guillardia theta (Cryptomonas phi)  
 OC Chloroplast.  
 CC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
 OX NCBI\_TaxID=55529;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 PC MEDLINE=92128221; PubMed=9223392;  
 RA Douglas S.E., Penny S.L.;  
 RA "The plastid genome of the cryptophyte alga, Guillardia theta:  
 RT complete sequence and conserved synteny groups confirm its common  
 RT ancestry with red algae."  
 RL J. Mol. Evol. 48:236-244(1999).  
 CC FUNCTION: ATPase REQUIRED FOR THE CORRECT PLACEMENT OF THE  
 CC DIVISION SITE (BY SIMILARITY).  
 CC - SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL, AF041468; AAC35621.1;  
 DR InterPro: IPR000707; ATPase\_Para.  
 DR Pfam: PF00991; Para; 1.  
 KW Cell division; Septation; ATP-binding; Chloroplast.  
 FT NP BIND 10 17  
 FT SEQUENCE 269 AA; 29455 MW; BC3C6B54E689A2 CRC64;

Query Match 42.0%; Score 694.5; DB 1, Length 269;  
 Best Local Similarity 55.9%; Pred. No. 1.9e-48;  
 Matches 146; Conservative 44; Mismatches 62; Indels 9; Gaps 3;

QY 59 RIVVITSGGKGGKTTTANVGLSLARYGSVAIDADLGRNDLLGLLENRVNTTVE 118  
 DB 3 RIVVITSGGKGGKTTTANVGLSLARYGSVAIDADLGRNDLLGLLENRVNTTVE 62  
 QY 119 VINGDCLDQALVDPKPMKSNFELLCTSKPPSKLPMWFGGKALEMVLNLTPTPEGSPDFI 178  
 DB 63 VINGDCLDQALVDPKPMKSNFELLCTSKPPSKLPMWFGGKALEMVLNLTPTPEGSPDFI 116  
 QY 179 IIDCPAGIDAGFTTATPANEAULTPTDITLRDADPVTSLEGTGIPDIKMTVNPVPT 238  
 DB 117 IIDCPAGIDAGFTTATPANEAULTPTDITLRDADPVTSLEGTGIPDIKMTVNPVPT 176  
 QY 239 DMKIGEDMMSVLDVQEMGLSLGVPEDESEVISTNPGPLVANKPPLAGLAEQAM 298  
 DB 177 DMKIGEDMMSVLDVQEMGLSLGVPEDESEVISTNPGPLVANKPPLAGLAEQAM 236  
 QY 299 PLVEQDSMAVWVEEPKRGPF...SFGG 319  
 DB 237 PLVEQDSMAVWVEEPKRGPF...SFGG 254

## RESULT 4

MIND\_SYN3

```

1D MIND SYNVS STANDARD; PRT; 266 AA.
AC 055900:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Septum site-determining protein mind (Cell division inhibitor mind).
OS MIND OR SL0289.
OS Synchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_Taxid=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.,
"Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. 1. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC DIVISION SITE. CELL DIVISION INHIBITORS MINC AND MIND ACT IN
CC CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE
CC POLAR Z RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE
CC CELL TO DESTABILIZE FT2Z FILAMENTS THAT HAVE FORMED BEFORE THEY
CC MATURE INTO POLAR Z RINGS (BY SIMILARITY).
CC -1- SUBUNIT: INTERACTS WITH MINC AND FT2Z (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D64005; BAA10662.1; -
DR InterPro; IPR000707; ATPase_Para.
DR InterPro; IPR000392; NitrogenaseII.
DR Pfam; PF00142; fer4_NiH; 1.
DR Pfam; PF00991; Para; 1.
KM Cell division; Septation; ATP-binding; Membrane; Complete proteome.
FT NP_BIND 10 17
FT ATP (POTENTIAL).
SQ SEQUENCE 266 AA; 29047 MW; 4BC37880F086C9BF CRC64;
Query Match 38.9%; Score 643; DB 1; Length 266;
Best Local Similarity 52.8%; Pred. No. 2,4e-44;
Matches 130; Conservative 51; Mismatches 59; Indels 6; Gaps 2;
OY 59 RIVVITSGKGVGKTTTANVGLSLARYGFSVVA:IDAD:G:RNLDLLGLENRVYTCVE 118
DB 3 RIVVITSGKGVGKTTTANVGLSLARYGKRVV:IDAD:G:RNLDLLGLENRVYTAID 62
OY 119 VINGOCDRLDQALVRDKWSNFFELCTSKRSTLPMGFGKAL:EWLVDALKTRPESGPDFI 178
DB 63 VLADECTIDKALVKKRLEPNLVLPAAQWRSK--DAINAEQMSQVEGDKX---FPIYI 116
OY 179 IIDCPAGIDAGFITATTANENAVLTTDPITLAPADPVYTGLEFDG:FDIKMYNRPFT 238
DB 117 IIDCPAGIDAGFRRNVAAPQEAIVTTPMSAVRADRIYGLLEAEDIGKSLYNRLRP 176
OY 239 DMKGEDMMSVLVDQVOMGLSLGVIPEDSEYIRSTNGRFP:LVNKPPPTLGLAFEGQAAW 298
DB 177 EMWQDNQMSVIEDIDLLAVPLIGLIPDQKIIISTNGEPLVMEKLSVGLAFQNIAR 236
OY 299 RIVEQD 304
DB 237 RLEGQD 242

```

```

1D MIND_BACSU STANDARD; PRT; 268 AA.
AC Q01464;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Septum site-determining protein mind (Cell division inhibitor mind).
OS MIND OR DIVVB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RA MEDLINE=93015732; PubMed=1400225;
RA Varley A.W., Stewart G.C.;
RT "The divVB region of the Bacillus subtilis chromosome encodes
RT homologs of Escherichia coli septum placement (mindC) and cell shape
RT (mreBD) determinants.";
RL J. Bacteriol. 174:6729-6742(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RA MEDLINE=93211302; PubMed=8459776;
RA Lee S., Price C.W.;
RT "The mind locus of Bacillus subtilis lacks the mind determinant that
RT provides topological specificity to cell division.";
RL Mol. Microbiol. 7:601-610(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93015731; PubMed=1400224;
RA Levin P.A., Margolis P.S., Setlow P., Losick R., Sun D.;
RT "Identification of Bacillus subtilis genes for septum placement and
RT shape determination.";
RL J. Bacteriol. 174:6717-6728(1992).
CC -1- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC DIVISION SITE. CELL DIVISION INHIBITORS MINC AND MIND ACT IN
CC CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE
CC POLAR Z RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE
CC CELL TO DESTABILIZE FT2Z FILAMENTS THAT HAVE FORMED BEFORE THEY
CC MATURE INTO POLAR Z RINGS (BY SIMILARITY).
CC -1- SUBUNIT: INTERACTS WITH MINC AND FT2Z (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -1- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95582; AAA22609.1; -
DR EMBL; Z15113; CAA78818.1; -
DR EMBL; M96343; AAA22401.1; -
DR EMBL; Z99118; CAA14759.1; -
DR PIR; S31205; S31205.
DR PIR; G45235; G45235.
DR PIR; P45240; P45240.
DR Subtilist; BG10330; mind.
DR InterPro; IPR000707; ATPase_Para.
DR Pfam; PF00991; Para; 1.
KM Cell division; Septation; ATP-binding; Membrane; Complete proteome.
FT NP_BIND 10 17
FT ATP (POTENTIAL).
SQ SEQUENCE 268 AA; 29407 MW; 6665B9F693F58A9B CRC64;
Query Match 31.6%; Score 523; DB 1; Length 268;
Best Local Similarity 41.1%; Pred. No. 9,4e-35;
Matches 111; Conservative 68; Mismatches 77; Indels 14; Gaps 7;
OY 61 VVITSGKGVGKTTTANVGLSLARYGFSVVA:IDAD:G:RNLDLLGLENRVYTCVEVI 120
DB 5 VVITSGKGVGKTTTANVGLSLARYGKRVV:IDAD:G:RNLDLLGLENRVYTVLVVV 64

```

RESULT 5  
MIND\_BACSU

```

QY 121 NDCRDLQALVDPKRWNSFELLCSIKPSKSLPMGFGKALEMLVDALKTRPGSPDFII 180
DB 65 ESRCKKHQALVMDKRPDDL-LYLMPAAGTSDKTAVAPEQIKMVGELKQ---EFDVYII 119
QY 181 DCPAGIDAGFITAIPANAVLVTPDITAPADAPRVGTLLEC-DGIRDIKINVRVATD 239
DB 120 DCPAGIEQGYKNAVSGADAVITVTPETISAVRDADRIICGLEEENIEPPRLVVRIRRH 179
QY 240 MIKGEDMSVLDPQEMTGLSLGVIPEDSEVIRSTRNGFPVLNRPPTLAGLAFQAAAR 299
DB 180 LKNGDITMDIDELVHLSIDLIGVANDDEVIKASHNCFIAMD PKRASIAYANINAR 238
QY 300 LVEQDSMKAVMWEEEPKRGFF---SFG 325
DB 239 LD-GESVPLQVLEQNK--GNMAKIKSPFG 265

RESULT 6
MIND_HELPY
ID MIND_HELPY STANDARD: PRT: 268 AA
AC 025098;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Septum site-determining protein mind (Cell division inhibitor mind).
OS Mind OR HPO331.
NC Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
PC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kurlavsky A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Richardson J., Zhou L., Kirkness E.F., Peterson S.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Kap P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC DIVISION SITE. CELL DIVISION INHIBITORS MIND AND MIND ACT IN
CC CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE
CC POLAR Z RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE
CC CELL TO DESTABILIZE FTSZ FILAMENTS THAT HAVE FORMED BEFORE THEY
CC MATURE INTO POLAR Z RINGS (BY SIMILARITY).
CC -1- SUBUNIT: INTERACTS WITH MIND AND FTSZ (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AE000551; AAD07400.1;
CC TIGR: HP0331;
CC InterPro: IPR000707; ATPase_Para.
CC InterPro: IPR000392; NitrogenaseII.
CC Pfam: PF00142; fer4_NiFH.1.
CC Cell division, Septation, ATP-binding, Membrane, Complete proteome.
CC NP_010170.17 ATP (POTENTIAL).

```

```

SQ SEQUENCE 268 AA; 29294 MW; 9E14766E8D30947E CPC64;
Query Match 30.6%; Score 506.5, DB 1, Length 268;
Best Local Similarity 40.1%; Pred. No. 1.9e-33;
Matches 110; Conservative 65; Mismatches 80; Indels 19; Gaps 5;

QY 60 IVVTSKGVKVTATTAVALVGLSLAPYFSSVAITCALCLPRLCLLGLPVPVNTGVVEV 119
DB 4 VVITSGKGVGKSTTTANLAIGLAEGRKVAVDPDILRNIDMLIGLNRIVDDVDV 63
QY 120 INDCRDLQALVDPKRWNSFELLCSIKPSKSLPMGFGKALEMLVDALKTRPGSPDFII 179
DB 64 MEKCNLSQALITDKTKNLSPLAASCKDKNIDKEKVAI-LINALP---ADFDYIL 117
QY 180 IDCPAGIDAGFITAIPANAVLVTPDITAPADAPRVGTLLEC-DGIRDIKINVRVATD 239
DB 118 IDSPAGIESGFHAILHADMALVVTPEVSLRSDSPVVGIIIDAKSNPAKGMVHKHLI 177
QY 233 VVRVPTMIKGEDMSVLDPQEMTGLSLGVIPEDSEVIRSTRNGFPVLNRPPTLAGLA 292
DB 178 INPLKPELVANGEMISIEVLTLCFLIGIIPEDHIIETATRGFFVI PTCGCAVA 235
QY 293 FEQAAARLVEQDSMKAVMWEEEPKRGFFSFG 326
DB 236 YQRTTRITGEE...VEYVEFAKRGFFSALKG 265

RESULT 7
MIND_HELPY
ID MIND_HELPY STANDARD: PRT: 268 AA.
AC 092MAR;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Septum site-determining protein mind (Cell division inhibitor mind).
GN Mind OR JHP0314.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
PC MEDLINE=99130557; PubMed=9923682;
RX Alm R.A., Ling L.S.L., Moir P.T., King R.L., Brown R.D., Draig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Urita-Nickelsen M., Mills D.W., Ives C.,
RA Gibson R., Weiberg D., Mills S.D., Jiang Q., Taylor P.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC DIVISION SITE. CELL DIVISION INHIBITORS MIND AND MIND ACT IN
CC CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE
CC POLAR Z RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE
CC CELL TO DESTABILIZE FTSZ FILAMENTS THAT HAVE FORMED BEFORE THEY
CC MATURE INTO POLAR Z RINGS (BY SIMILARITY).
CC -1- SUBUNIT: INTERACTS WITH MIND AND FTSZ (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AE001468; AAD05905.1;
CC InterPro: IPR000707; ATPase_Para.
CC InterPro: IPR000392; NitrogenaseII.
CC Pfam: PF00142; fer4_NiFH.1.

```





DR Ecogene; EGI0597; mind.  
 DR InterPro; IPR00707; ATPase\_Para.  
 DR InterPro; IPR00348; Arsa\_ATPase.  
 DR Pfam; PF00991; Para; 1.  
 DR Pfam; PF02374; Arsa\_ATPase; 1.  
 DR Cell division; Septation; ATP-binding; Inner membrane;  
 KW Complete proteome.  
 FT INIT\_MET 0  
 FT NP\_BIND 9 16 ATP (POTENTIAL).  
 FT MUTAGEN 14 14 G->S: LESS EFFECTIVE THEN WILD-TYPE.  
 FT MUTAGEN 15 15 K->Q: LOSS OF ACTIVITY.  
 FT MUTAGEN 15 16 KT->QP: LOSS OF ACTIVITY.  
 SO SEQUENCE 269 AA, 29483 MW, C21B2A476A190B1 CRC64.  
 Query Match 29.0%; Score 473; DB 1; Length 269;  
 Best Local Similarity 39.9%; Pred. No. 3, 1e-31;  
 Matches 113; Conservative 58; Mismatches 82; Indels 30; Gaps 8;

QY 59 RIVVITSGKGVGKTTTANVGLSLAPYGSVAITADLGRN.DLLGLENPNVTCVE 118  
 DB 2 RIVVITSGKGVGKTTTANVGLSLAPYGSVAITADLGRN.DLLGLENPNVTCVE 118  
 QY 119 VINGCDRLDQALVPPKPSNPFLLCTSKPPSKLPWGFAGALEWLVDALEKTPPGSPDF 178  
 DB 62 VIOGDAITNQAALIKDKPTENLYILPASQTRDQALTRGVAK--VLDDLKAM--DPEFI 116  
 QY 179 IIDCPAGIDAGFTTATPANENAVLTPDITAPADPVTGLEC-----DGIPTDY-- 230  
 DB 117 VCDSPAGIETGALMALYFDEAIITTNPEVSSVPDSDPLIGILSKSPRAENGEPIKEH 176  
 QY 231 MIVNRRTDMIGEDMMSVLDVQEMGLGVIPEDSEVIRSTNRGFLVANKPPTLAG 290  
 DB 177 LLITPRNPSPVSRGMLSMEDVLEIRIKLVVIRPDQSVLPANQCEPVLIDINAD-AG 235  
 QY 291 LAPEQAMPVLEQNSMAYVVEEP...KPGPFS-PFGS 326  
 DB 236 KAY-ADVEPILCEEPFPPEPEEKGFELKPLFG 269

RESULT 9  
 MIND\_THEMA STANDARD; PRT; 271 AA.  
 ID MIND\_THEMA STANDARD; PRT; 271 AA.  
 AC Q9X213;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, last sequence update)  
 DT 16-OCT-2001 (Rel. 40, last annotation update)  
 DE Septum site-determining protein mind (Cell division inhibitor mind).  
 GN MIND OR TM1870.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;  
 OC Thermotogaceae; Thermotoga.  
 OX NCBI\_TaxID=23336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MS8 / DSM 3109;  
 RX MEDLINE=9287316; PubMed=10360571;  
 RA Nelson K E, Clayton P A, Gill S P, Gwin M L, Dodson P J, Nelson J D, Peterson J D, Nelson M C, Ketchum K A, Haft D H, Hickey E K, Peterson J D, Nelson M C, Ketchum K A, McDonald L, Utechtack T F, Malek J A, Linher K D, Garrett M M, Stewart A M, Cotton M D, Pratt M S, Phillips C A, Richardson D, Heideberg J, Sutton G G, Fleischmann R D, Eisen J A, White O, Salzberg S L, Smith H O, Venter J C, Fraser C M;  
 RA "Evidence for lateral gene transfer between Archaea and Bacteria from RT genome sequence of Thermotoga maritima";  
 RL Nature 399:323-329(1999).  
 CC -1- FUNCTION: ATPase REQUIRED FOR THE CORRECT PLACEMENT OF THE DIVISION SITE. CELL DIVISION INHIBITORS MIND AND MIND ACT IN CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE POLAR 2 RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE CELL TO DESTABILIZE FTSZ FILAMENTS THAT HAVE FORMED BEFORE THEY MATURE INTO POLAR 2 RINGS (BY SIMILARITY).  
 CC -1- SUBUNIT: INTERACTS WITH MING AND FTSZ (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (by similarity).

CC -1- SIMILARITY: BELONGS TO THE PAPA FAMILY MIND SUPERFAMILY  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-stb.ch/announce/> or send an email to [license@isb-stb.ch](mailto:license@isb-stb.ch)).  
 CC EMBL; AE001824; AAC36932.1; -  
 CC TRIP; TM1870; -  
 DR InterPro; IPR00707; ATPase\_Para.  
 DR Pfam; PF00991; Para; 1.  
 DR Cell division; Septation; ATP-binding; Membrane; Complete proteome.  
 FT NP\_BIND 10 17 ATP (POTENTIAL).  
 SO SEQUENCE 271 AA, 29483 MW, EC7ED3F15D64MCA CRC64.  
 Query Match 29.7%; Score 475.5; DB 1; Length 271;  
 Best Local Similarity 40.9%; Pred. No. 5, 9e-32;  
 Matches 105; Conservative 47; Mismatches 68; Indels 37; Gaps 5;

QY 60 IIVITSGKGVGKTTTANVGLSLAPYGSVAITADLGRN.DLLGLENPNVTCVE 119  
 DB 4 IIVITSGKGVGKTTTANVGLSLAPYGSVAITADLGRN.DLLGLENPNVTCVE 119  
 QY 120 INGCPLDQALVPPKPSNPFLLCTSKPPSKLPWGFAGALEWLVDALEKTPPGSP-- 175  
 DB 64 VNGKSPQDALVHKMLKNLYL--PASQ-----IATKEMISNDMK 103  
 QY 176 -----DEIIDCPAGIDAGFTTATPANENAVLTPDITAPADPVTGLEC 225  
 DB 104 AIVKELIHFPYIITIDSPAGIERGFNAVAPPELVVTPPEIPALSTADPVGILENIG 163  
 QY 226 IPD--IPWIVPVTWVTPKPFMMSVLQVQEMGLGVIPEDSEVIRSTNRGFLVANK 283  
 DB 164 FSDEKIVHIVIRFPHVVKCEMLTTDDIPIHTLLEITAVIPSEEDIVASNTQIFVSIN 223  
 QY 284 KPPLTAGLAPEQAMRL 300  
 DB 224 GNSRIS KPENLARI 239

RESULT 10  
 MIND\_BUCAI STANDARD; PRT; 269 AA.  
 ID MIND\_BUCAI STANDARD; PRT; 269 AA.  
 AC P57411;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, last sequence update)  
 DT 16 OCT-2001 (Rel. 40, last annotation update)  
 DE Septum site-determining protein mind (Cell division inhibitor mind).  
 GN MIND OR BU326.  
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 OS symbiotic bacterium).  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=118099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Tokyo 1998;  
 RX MEDLINE=20445173; PubMed=10993077;  
 RA Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H;  
 RA "Genome sequence of the endocellular bacterial symbiont of aphids RT Buchnera sp. Aps.";  
 RL Nature 407:81-86(2000).  
 CC -1- FUNCTION: ATPase REQUIRED FOR THE CORRECT PLACEMENT OF THE DIVISION SITE. CELL DIVISION INHIBITORS MIND AND MIND ACT IN CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE POLAR 2 RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE CELL TO DESTABILIZE FTSZ FILAMENTS THAT HAVE FORMED BEFORE THEY MATURE INTO POLAR 2 RINGS (BY SIMILARITY).  
 CC -1- SUBUNIT: INTERACTS WITH MING AND FTSZ (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE PAPA FAMILY MIND SUPERFAMILY

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, AP01119, BAB13034.1, -  
 CC InterPro: IPR000707; ATPase\_Para.  
 CC InterPro: IPR003348; Arsa\_ATPase.  
 CC Pfam: PF00991; Para; 1.  
 CC Pfam: PF02374; Arsa\_ATPase; 1.  
 CC Cell division, Septation; ATP-binding; Membrane; Complete proteome.  
 CC INIT MET 0 BY SIMILARITY.  
 CC NP\_BIND 9 16 ATP (POTENTIAL)  
 CC SEQUENCE 269 AA; 29805 MW; 4FFD5764549C8B27 CRC64;

Query Match 27.8%; Score 460; DB 1; Length 269;  
 Best Local Similarity 38.3%; Pred. No. 1e-29;  
 Matches 106; Conservative 60; Mismatches 91; Indels 20; Gaps 5;

QY 59 RIVVITSGKGVGKTTTANVGLSLARYGFSVAIDADLGLRNDLLGLENVYTCVEI 118  
 DB 2 RIVVITSGKGVGKTTSSAIGTGLAQGKKTIVDPDGLRNDLLMGCRRTVYDFIN 61  
 QY 119 VINGCRLDQALVPRKPSNFELLCISKPPSKLPMWFGG--KALEWLVDAKTRPESGP 176  
 DB 62 VVGQDATTALQAIKKKKTNNLFLPASQTRKDALTRIGVEKVTETI-----KNKFD 114  
 QY 177 FLIIDPAGIDAGFTTATTPANEAVLVTPDITALRDADRVGTGLCEGGRIDK----- 230  
 DB 115 FLICSPAGIEGALALYFADEAITTNPEVSSVRDSDRLGLGISKSKRAENKIPK 174  
 QY 231 --MIVNRVTDMIKEDNMVSLDVQEMVGLSLGVIPDSEVIRSTNPGPPLVLPPTL 288  
 DB 175 EXLLTRVNPVPRKKGEMLSMTDVLDIQIPILGIVPEDQSLVASHQGEIILIDINGN 233  
 QY 289 AGLAFEGQAMRIVEODSKMAMVVEEPRKRGFFSFFG 325  
 DB 234 AGCAVSTVNRLLGEE--RHFRFIEEKKSFLRLFG 268

RESULT 11  
 Y547 METUA  
 ID Y547 METUA STANDARD: PRT; 264 AA.  
 AC 0579677  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical ATP-binding protein Mj0547.  
 GN Mj0547.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerevage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,  
 RA Uitterlbeck T.P., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii.";  
 RT Science 273:1058-1073 (1996).  
 RL Science 273:1058-1073 (1996).  
 CC -!- SIMILARITY: HIGH, TO M.JANNASCHII Mj0169 AND Mj0410.

CC -----  
 CC -!- SIMILARITY: SOME, TO PROKARYOTIC MIND AND TO THE MRP FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, U67504; AAB98539.1, -  
 CC TIGR: Mj0547.  
 CC InterPro: IPR000707; ATPase\_Para.  
 CC Pfam: PF00991; Para; 1.  
 CC Hypothetical protein; ATP-binding; Complete proteome  
 CC NP\_BIND 13 21 ATP (POTENTIAL).  
 CC SEQUENCE 264 AA; 27737 MW; B9AF5108CD0DC63 CRC64;

Query Match 18.1%; Score 300; DB 1; Length 264;  
 Best Local Similarity 30.5%; Pred. No. 6e-17;  
 Matches 82; Conservative 57; Mismatches 104; Indels 26; Gaps 6;

QY 61 VITSGKGVGKTTTANVGLSLARYGFSVAIDADLGLRNDLLGLENVYTCVEI 120  
 DB 9 IAIISGKGTGKTITISANLAVALKRKYAVLDADIMANLELNGLSK-----PVT 62  
 QY 121 NGDCRLDQALVPRKPSNFELLCISKPPSKLPMWFGG--FGALEWLVDAKTRPESGP 175  
 DB 63 LNDVLAGADIKDAIYEGPEGVIV-----IPAGVSLERFRRAKPEKLEVKAIHD-LV 115  
 QY 176 FLIIDPAGIDAGFTTATTPANEAVLVTPDITALRDADRVGTGLCEGGRIDK----- 235  
 DB 116 FLIIDPAGIGKTTTALISADGLVWVPELSSISDAKTIATIKRSTDIIGAVNR 175  
 QY 236 VRTDMIKGEDNMVSLDVQEMVGLSLGVIPDSEVIRSTNPGPPLVLPPTL 295  
 DB 176 VSNESTE---LGKALEITLLEVVICVPEDEHYKKAAGFIPVIMFPDSPAQAIN 231  
 QY 296 AAMRLVEDSKMAMVVEEPRKRGFFSFF 324  
 DB 232 IAAFLIG----AFYEAQLKKKESFLSKF 256

RESULT 12  
 Y169 METUA  
 ID Y169 METUA STANDARD: PRT; 263 AA.  
 AC 057633;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical ATP-binding protein Mj0169.  
 GN Mj0169.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerevage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,  
 RA Uitterlbeck T.P., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii.";  
 RT Science 273:1058-1073 (1996).  
 RL Science 273:1058-1073 (1996).  
 CC -!- SIMILARITY: HIGH, TO M.JANNASCHII Mj0547, SOME, TO Mj0410 AND  
 TO Mj0924.

```

CC - SIMILARITY. SOME TO PROKARYOTIC MIND AND TO THE MRP FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U67474; AAB98154.1;
CC
CC TIGR: M20169;
CC
CC InterPro: IPR000707; ATPase_Para
CC Pfam: PF00991; Para; 1
CC Hypothetical protein; ATP-binding; Complete proteome.
CC NP BIND 10 17 ATP (POTENTIAL).
CC
CC SEQUENCE 263 AA; 28050 MW; 292F8B38E72AC94 CRC64;
CC
Query Match 14.6%; Score 242; DB 1; Length 263.
Best Local Similarity 29.3%; Pred. No. 2 5e-12;
Matches 77; Conservative 59; Mismatches 99; Indels 28; Gaps 9;
CC
OY 60 IVVITSGKGVGKTTTNAVGLSLARYGSVAIDADLGLNLDLLGLNRYNYTCV 119
DB 5 ITTASGKGVGKTTTSLAVALKKGLKVALIDGDI.SMANLGLFMEXK-KPSLHEV 63
OY 120 INGDCLDQALYRD--KMSNFELLCISPRSKLPMFGSKALEVLVDALKTRPEGSP 176
DB 64 LS-----EADVDRAIYKIKITGVVL---PTSLSLBEGYKKSDIDLDPDVNEVADF-D 113
OY 177 FTIIDCPAGIDAGFTIATPANEAVALVTPDITLRDADPFTGLFCGJIDIMIVNPV 236
DB 114 VVITAPALNEMATHALAKKLLVTPPEFSTIIDAVRLKESANAGTLMGVLTNRV 173
OY 237 PTTMTVGERMMSVLQVCEMLGSLGVIPEDSEVIRSTNCPFLV...LNKPPTLAGLAF 293
DB 174 GRDF--GE--MGPDEIEMLIKGVLEVPEDENVSAALYKMSVLEYEYKNSPAS... 223
OY 294 EOAAMRLVEQDSMKAMVNEEPK 316
DB 224 -QAVMKLASIIAGVPIYEDIK 245
CC
RESULT 1
YLXH BACSU STANDARD; PRT; 298 AA
AC P40742;
DT 01-FEB-1995 (rel. 31, Created)
DT 01-FEB-1995 (rel. 31, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Hypothetical protein ylxH.
GN YLXH.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCB TaxID=1423;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE=95078462; PubMed=7987014;
RA Kirsch M L, Carpenter P R, Ordal G W;
RA "A putative ATP-binding protein from the che/fla locus of Bacillus
RA subtilis";
RA Subtilis";
RA DNA Seq. 4:271-275(1994)
RN 12
RP SEQUENCE FROM N.A
RA STRAIN=168;
RA MEDLINE=98044073; PubMed=9184377;
RA Kunze F, Ogasawara N, Moszer I, Albertini A.M., Alloni G,
RA Azevedo V, Barteto M G, Bessieres P, Bolotin A, Borcher S,
RA Boriss P, Boursier L, Brans A, Braun M, Brigelli S.C., Bron S,
RA Brouillet S, Brunsch C V, Caldwell B, Capuano V, Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehlich S.D., Emerson P.T.,
RA Ertlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

```

```

RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
RA Ghm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy R.J., Haga K., Haejch J., Harwood C.P., Henaut A.,
RA Hilbert H., Holappel S., Hosono S., Hullo M.F., Ilaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kjaer-Rianhard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kunita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Maeda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Potwolik S., Prescott A.M.,
RA Prescan E., Puic P., Punnelle B., Rapoport G., Pay M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scorfione F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpsita P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wandut P., Wedler E., Wedler H., Weitzenecker T.,
RA Winfers P., Wipar A., Yamamoto R., Yaneva K., Yastmotic K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RA "The complete genome sequence of the Gram-positive bacterium Bacillus
RA subtilis";
RA Nature 390:249-256(1997)
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X73682; CAAC52034.1;
CC EMBL: Z99112; CAB1514.1;
CC PIR: S34194; S34194.
CC Subtilis; BG10254; YLXH.
CC InterPro: IPR000707; ATPase_Para.
CC Pfam: PF00991; Para; 1
CC Hypothetical protein; ATP-binding; Complete proteome.
CC NP BIND 39 46 ATP (POTENTIAL).
CC
CC SEQUENCE 298 AA; 33160 MW; AF6FE71FA31C972 CRC64;
CC
Query Match 12.9%; Score 214; DB 1; Length 298.
Best Local Similarity 25.2%; Pred. No. 5.1e-10;
Matches 75; Conservative 56; Mismatches 119; Indels 50; Gaps 10;
CC
OY 32 VNPSPRSPIRSVLQFNK--PELAGETPRIVITSGKGVGKTTTNAVGLSLARYGS 89
DB 3 NMRVQAAATLAKKKEKEPVLPMVYSQAKTLAVISGKGVGKSNITLMNALALQKGR 62
OY 90 VVAIDADLGLNLDLLGLNRYNYTCVINGDCRLDQALYVDKXMSNFELLCISPRPS 149
DB 63 VLLIDLDIGKMNIDILIG--NSSATIIDVLTDRKPLQSL-----S 102
OY 150 KLPWGF---GKALE-----WLVDAKLTREPGSP--DFIIDCPAGIDAGFTIAT 195
DB 103 VQPKLRFYISGATGLDWMFQCDQKMTFFANLSHALSCPDVYLFPMGAGLSQDQPLIL 162
OY 196 FANEAVALVTPDITLRDPA-CPVTGLLETCGIDPIYVNPVFTWIKEDMS----- 248
DB 163 SAEDILITTEPFAIMDAVSAVHLVLTENKLSMKVAVAPCP-PQVEGIDAPAPISPT 221
OY 249 --VLDQVCEMLGSLGVIPEDSEVIRSTNCPFLV...LNKPPTLAGLAFQAMVLTQDSM 306
DB 222 HPELVQ---VQAGVSDVIVSAVVEQVPEFPIYTCAKKRCVRIILADALFEET 277
OY 307 K 307
DB 278 R 278
CC
RESULT 14

```

ID	MRP_SYNY3	STANDARD:	PRT:	353 AA.
AC	MRP_SYNY3			
AC	P53583			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Mrp protein homolog.			
GN	MRP OR SLP067			
OS	Synechocystis sp. (strain PCC 6803).			
OC	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.			
OK	NCBI TaxID=1148;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
PX	MEDLINE=96127529; PubMed=8590279;			
RA	Kaneke T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,			
RA	Sugitara M., Tabita S.,			
RT	"Sequence analysis of the genome of the unicellular cyanobacterium			
RT	Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb			
RT	region from map positions 64k to 92k of the genome."			
RL	DNA Res. 2:153-166(1995).			
CC	-1- FUNCTION: NOT KNOWN.			
CC	-1- SIMILARITY: BELONGS TO THE MRP/NBP35 FAMILY OF ATP-BINDING			
CC	PROTEINS.			
CC	-----			
CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its use by			
CC	non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; D64001; BAA10297.1; -			
DR	InterPro; IPR002744; MRP_F59.			
DR	InterPro; IPR000808; MRP_family.			
DR	Pfam; PF01883; DUF59.1.			
DR	ProDom; PD005595; DUF59.1.			
DR	PROSITE; PS01215; MRP.1.			
KW	ATP-binding; Complete proteome.			
FT	NP BIND 105 112 ATP (POTENTIAL).			
EQ	SEQUENCE 353 AA; 37131 MW; 891DCOE247113841 CRC64;			
Query Match	12.9%; Score 213; DB 1; Length 353;			
Best Local Similarity	26.1%; Pred. No. 7.7e-10;			
Matches	82; Conservative 55; Mismatches 115; Indels 62; Gaps 12;			
OY	23 KTLSSPPFVNPNPSRPPIPSVILQFNRPKPELAGEPPIPLVITISGKGVCGKTTTAYVGS 82			
Db	76 KVTAEPTQCKSLSDPDRQ-----VGYNINIAISSGKGVGISTAVANVAVA 121			
OY	83 LARGFSSVAIDADLGLRNLIDLISLENVNTYCYEVINSDCLDLQALVDEKZMSFELL 142			
Db	122 LAQGAAGVGLIDLDIVGPNAPTMGLSG-----AAVOQNS-----PQGEVLEPVF-NHGIR 172			
OY	143 CISKPRSKLPMGF-----GGALFLMWLDALKTRPE-GSPDPILIIDCPGI-DAGE 190			
Db	173 MVS-----MGFLIDPDQVIMRGMGLGIIIPQLLYGVNMGALDYLILVMPSPYSDAQ 225			
OY	191 -ITAITPANEAVLVTPPDITALEKADQVTTGLDCTGIPDITMIVNP---VRTMIK.... 242			
Db	226 TLTOSVPMAGAVIVTTPTQTSILIDAPPG; KMPQGVGVNVVLGIVENMEYFIPPLPLPQYD 285			
OY	243 -----GEDMWSVLVDQEMGLSLGIVPEDEBVARSTNRGFPVLVLRKPTLAGLAEQA 296			
Db	286 LFGSGGGEKRSKEINV-----PLIGCVLELIGLEEGSDKGVPIVVSQPSASAKALTAI 339			
OY	297 AMLRVEDQSMKAYM 310			
Db	340 AKQIAGKVSMAALV 353			

```

AD MRP_AQUAE STANDARD: PRT: 364 AA.
IC 066946:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MRP protein homolog.
GN MRP OR AQ.737.
OS Aquifex aeolicus.
OC Bacteriia, Aquificae, Aquificae (class); Aquificales, Aquificaceae;
OC Aquifex.
OC NCBI_taxid=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RA "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RT Nature 392:353-358 (1998).
RL -!- FUNCTION: NOT KNOWN.
CC -!- SIMILARITY: BELONGS TO THE MRP/NEP35 FAMILY OF ATP BINDING
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000705; AAC06935.1; -
DR InterPro: IPR002744; DUF59.
DR InterPro: IPR000808; MRP_family.
DR Pfam: PF01883; DUF59; 1.
DR PROSITE: PS01215; MRP; 1.
KW ATP-binding; Complete proteome.
KW NP BIND 120 127 ATP (POTENTIAL).
SQ SEQUENCE 364 AA, 39492 MW, 330AF95C19.A272A CRC64;

Query Match 12.8%; Score 211; DB 1; Length 364;
Best Local Similarity 27.1%; Pred. No 1,2e-09;
Matches 75; Conservative 50; Mismatches 110; Indels 42; Gaps 10;

OY 35 PSRRSPFSLQFNRPFLASGTEPRIVITSSKSGVSKTTTANVGLSLARGFVSVAID 94
DB 96 PQTQDEP-----FTRK-KVPG-VKHIIVAGSGKSGVSKSTVAANLVAALSDGKKGILD 148
OY 95 ADIGLRRLDILGLENRVVYTCVEYINSGCFILCALVDFRPMNSFELLCSIPRSKLPWG 154
DB 149 ADVYGSVPLTGLGK-----EAVYVD-QFCIIIVEYKGIKILSIGMPMSEDTPII 200
OY 155 FGKAL-EMVLDAKTRPGSPDPIIIDCPAGIDAGTIAI--TPANEAVLTPDITAI 211
DB 201 WRGPMKMLATELSTSKGNLDFVMDLPSTEDVDITLAQNVVELTGAVVVTTPODVAL 260
OY 212 RQDPRVATGLECDGSRDIKMYINRFTMI-----GCEDMASVLEQNTL 256
DB 261 ADVKAVSM-----FREVVIPVLGIENMAVFCPSDKQRYVIFGKQ---VAEPANAY 311
OY 257 GLSLGVIPEDSEVIRSTNRGFFVLNKPPLAGIAF 293
DB 312 GLKILGISIPIDEVAEKSDKGPVIYISHPSEVAKAF 348

```



GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: April 16, 2003, 09:04:57 ; Search time 15 Seconds  
(Without alignments) 639,458 Million cell updates/sec

Title: US-10-067-989-1  
Perfect score: 1654  
Sequence: 1 MASH/PTSTHQSUIIPSSL KATWVEEPKPPFPFPFG 226

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1 pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	144	8	7	US-09-134-001C-4842	Sequence 4842, Ap
2	88.5	5.4	264	US-08-969-644-16	Sequence 16, Appl
3	88.5	5.4	264	US-08-444-189-16	Sequence 16, Appl
4	88.5	5.4	264	US-08-468-544-16	Sequence 16, Appl
5	88.5	5.4	519	US-08-997-4450-2	Sequence 2, Appl
6	88	5.3	376	US-09-461-474-6	Sequence 6, Appl
7	88	5.3	732	PCT-US95-17026-2	Sequence 2, Appl
8	87	5.3	249	US-08-597-235-5	Sequence 5, Appl
9	87	5.3	249	US-08-746-682A-5	Sequence 5, Appl
10	86.5	5.2	416	US-09-320-878-18	Sequence 18, Appl
11	86.5	5.2	416	US-09-105-537-39	Sequence 39, Appl
12	86	5.2	410	US-09-134-001C-3927	Sequence 3927, Ap
13	85.5	5.2	1068	US-08-085-1898-11	Sequence 11, Appl
14	83.5	5.0	3077	US-08-416-603-4	Sequence 4, Appl
15	83.5	5.0	3457	US-08-896-320-1	Sequence 1, Appl
16	82.5	5.0	367	US-08-896-320-3	Sequence 3, Appl
17	82.5	5.0	367	US-08-896-320-3	Sequence 3, Appl
18	82.5	5.0	396	US-09-461-474-2	Sequence 2, Appl
19	82.5	5.0	670	US-08-980-080-2	Sequence 2, Appl
20	82.5	5.0	756	US-08-434-730-16	Sequence 16, Appl
21	82	5.0	258	US-09-504-358-12	Sequence 12, Appl
22	82	5.0	258	US-09-504-358-12	Sequence 12, Appl
23	81.5	4.9	5087	US-09-144-085-1	Sequence 1, Appl
24	81	4.9	385	US-09-570-778A-12	Sequence 12, Appl
25	81	4.9	920	US-08-930-966A-8	Sequence 8, Appl
26	81	4.9	1705	US-08-669-785-4	Sequence 4, Appl
27	80.5	4.9	332	US-09-172-952-11	Sequence 11, Appl

#### ALIGNMENTS

28 80.5 4.9 812 1 US-08-446-794A-2 Sequence 2, Appl  
29 80.5 4.9 812 1 US-08-750-007-3 Sequence 3, Appl  
30 80.5 4.9 812 2 US-08-945-024-2 Sequence 2, Appl  
31 79 4.8 1041 1 US-08-220-151-4 Sequence 4, Appl  
32 79 4.8 1041 1 US-08-413-118-4 Sequence 4, Appl  
33 79 4.8 1041 3 US-08-473-446-4 Sequence 4, Appl  
34 78.5 4.7 371 2 US-08-837-593-8 Sequence 8, Appl  
35 78.5 4.7 676 4 US-09-085-1998-7 Sequence 7, Appl  
36 78.5 4.7 1166 4 US-09-134-001C-4522 Sequence 4522, Ap  
37 78 4.7 332 2 US-09-461-474-15 Sequence 15, Appl  
38 78 4.7 332 4 US-08-674-149A-2 Sequence 2, Appl  
39 78 4.7 937 1 US-08-253-155A-31 Sequence 31, Appl  
40 78 4.7 3567 2 US-07-642-734C-4 Sequence 4, Appl  
41 78 4.7 3567 3 US-08-433-032A-4 Sequence 4, Appl  
42 77.5 4.7 325 4 US-09-088-435-1 Sequence 1, Appl  
43 77.5 4.7 338 4 US-09-134-001C-4795 Sequence 4795, Ap  
44 77.5 4.7 2037 4 US-09-106-998-3 Sequence 3, Appl  
45 77 4.7 420 4 US-09-066-047-8 Sequence 8, Appl

RESULT 1  
US-09-134-001C-4842  
Sequence 4842, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
PRIOR FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 4842  
LENGTH: 359  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4842

Query Match 8.7%; Score 144; DB 4; Length 359;  
Best Local Similarity 24.3%; Pred. No. 9.3e-08;  
Matches 55; Conservative 44; Mismatches 79; Indels 48; Gaps 8;

QY 16 LPSSLSQKTLISSRFNNPSRR-SPIRSVLQFNRPDLAAGTRPRIVITSGKGVGKTT 74  
DB 85 LPSEVE-----RYIGKSEKPTTEELSQNNPVE-----FISIAS3P3SV3KST 130  
QY 75 TTAVNGSLARVGSVVAIDADLGRNLDLGLLENRVNVTCEVI---NDCRDLQAL 130  
DB 131 VAVNVAVALAEKGKVLVDADYIGFSVPDMGIDRPGDKEIIPVRHGVKYSMAF 190  
QY 131 VRDKRMSFELICTSKRSKLEPMGFGKALEMLVDALKTRPE-GSPDFITIDCPASID-- 187  
DB 131 FVEE-----NAFVWRGFMGKMLTFTEVQWGEIDVLLDLPRTGIV 235  
QY 188 AGFTAITPANEAVLVTPDITLADADR-----VNGLE 222  
DB 236 ALDVHSMIPSKKEIVTTPHTPAFAVAARAGAMAKTEHTILGIVR 281

RESULT 2  
US-08-969-644-16  
Sequence 16, Application US/08969644  
Patent No. 6096519  
GENERAL INFORMATION:  
APPLICANT: Ratti, Giulio  
APPLICANT: Comanducci, Maurizio



APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM: RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/969,644  
FILING DATE: 13-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/467,152  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE:  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-969-644-16

Query Match 5.4%; Score 88.5; DB 3; Length 264;  
Best Local Similarity 24.0%; Pred. No. 0.11; Indels 35; Gaps 7;  
Matches 44; Conservative 26; Mismatches 78;

61 VVTSKGVGKVTATTAVNGLSLARY GFSVVAIPADLGLRLDILLGLEHNVVTCVEV 119  
4 LVFCSTKGTGTTLSLVNVCNLAQFLGKVLADLD-PQSNLSSTG----- 50

120 INGDPLDQALVRKWSNFEL--LCISKPPS--KLPMGFQKALEWLVDAKTPREGS 174  
51 --ASVPSDQKGLHDIIVTSNDLSIICETKSDVDLIPASFSSSQFREL--DIHGFSSNN 106

175 -----PPFIIDCPAGIDAGFITAITPANEAVLVTTPDITALRADAVTGLLE 222  
107 LKLFLNEYCAPFYDICIIDTPPSLGGITKEAFVAGDKLIACLTPEPFSILGQKIREFLS 166

223 CDG 225  
167 SVG 169

RESULT 3  
US-08-444-189-16  
Sequence 16, Application US/08444189  
Patent No. 6110705  
GENERAL INFORMATION:  
APPLICANT: Ratti, Giulio

APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM: RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,189  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/180,528  
FILING DATE:  
APPLICATION NUMBER: US/07/991,512  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE:  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-444-189-16

Query Match 5.4%; Score 88.5; DB 3; Length 264;  
Best Local Similarity 24.0%; Pred. No. 0.11; Indels 35; Gaps 7;  
Matches 44; Conservative 26; Mismatches 78;

61 VVTSKGVGKVTATTAVNGLSLARY GFSVVAIPADLGLRLDILLGLEHNVVTCVEV 119  
4 LVFCSTKGTGTTLSLVNVCNLAQFLGKVLADLD-PQSNLSSTG----- 50

120 INGDPLDQALVRKWSNFEL--LCISKPPS--KLPMGFQKALEWLVDAKTPREGS 174  
51 --ASVPSDQKGLHDIIVTSNDLSIICETKSDVDLIPASFSSSQFREL--DIHGFSSNN 106

175 -----PPFIIDCPAGIDAGFITAITPANEAVLVTTPDITALRADAVTGLLE 222  
107 LKLFLNEYCAPFYDICIIDTPPSLGGITKEAFVAGDKLIACLTPEPFSILGQKIREFLS 166

223 CDG 225  
167 SVG 169

RESULT 4  
US-08-468-544-16  
Sequence 16, Application US/08468544

```
Patent No. 6248563
GENERAL INFORMATION:
APPLICANT: Ratti, Giulio
APPLICANT: Comanducci, Maurizio
APPLICANT: Tecce, Mario F.
APPLICANT: Giuliani, Marzia M.
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY
TITLE OF INVENTION: THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID
NUMBER OF SEQUENCES: 23
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 301 N. Washington Street
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22046-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent'n Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,544
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/661,820
FILING DATE: 28-FEB-1991
APPLICATION NUMBER: IT MI 91A000314
FILING DATE: 07-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Svensson, Leonard R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 1267.202P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-241-1300
TELEFAX: 703-241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-544-16

Query Match 5.4%; Score 88.5; DB 4; Length 264;
Best Local Similarity 24.0%; Pred No 0.11;
Matches 44; Conservative 26; Mismatches 78; Indels 35; Gaps 7;

QY 61 VITGKGVGKTTTANVGLSLARY-GFSVAIDADLGLRLDLLGLENNVYTCVEV 119
DB 4 LVFCSEKGTGKTTLSLVGNCGLAFLGKVLADLD-POSNLSGGG----- 50
QY 120 INGDRLDQALVPRGWSNFEI--LCISKPS--KLPMGFGKALEMLVDALKTPESS 174
DB 51 --ASVRSQKGLHDIVYSNDLSKISICTKDSVDLLPASFSSEQFREL--DIHRGSNN 106
QY 175 -----PDFIIDCPAGIDAGFTTAITPANEAUVLTTPDITALPDADPVTGLLE 222
DB 107 LKLFLNEYCAPFYDICIIDTPPSLGLTKFAFVAGDKLINCTLEPSPSIIGLQKIRFELS 166
QY 223 CDG 225
DB 167 SVG 169

RESULT 5
US-08-997-445D-2
Sequence 2, Application US/08997445D
Patent No. 6043342
```

```
GENERAL INFORMATION:
APPLICANT: Kocher, Olivier N.
TITLE OF INVENTION: PDZKI Protein Containing PDZ
TITLE OF INVENTION: Interaction Domains
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: David Prashker, Esq.
STREET: P.O. Box 67
CITY: Brookline
STATE: Massachusetts
COUNTRY: USA
ZIP: 02146
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.40 MB storage
COMPUTER: IBM PS/1
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,445D
FILING DATE: December 23, 1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: David Prashker
REGISTRATION NUMBER: 29,693
REFERENCE/DOCKET NUMBER: BIS-037
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 232-7509
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-445D-2

Query Match 5.4%; Score 88.5; DB 3; Length 519;
Best Local Similarity 19.3%; Pred No 0.33;
Matches 53; Conservative 38; Mismatches 76; Indels 107; Gaps 12;

QY 32 VNPSRSPRISRVLOFNRP--ELAGETPRIVITSGKGWKTITTTANVGLSLARYGF 88
DB 232 VDKETDKRHVQKIQFKRETASLKLPHQPRIVEMKKSNG-----YGF 255
QY 89 SVVA-----IDA-----DLGLRLDLLGLENNVYTCVEVINGDRLD----- 127
DB 256 YLRGSEKQGIINDIDSGSPAEAGLKNNDLVVA-----VNGSEVETLDDHDSVEMIRKG 311
QY 128 ----QALVRDPRGWSNFEI/LCISKP-----PSKLPMGFGKALEMLVDALKTPESSPPI 179
DB 312 GQCSLIVVCKETKMTLRAHFSFPIIVYQSELEFG-----SVKEAFATFTSL 360
QY 179 IIDCP-----AGIDAGFTTAITPANEAUVLTTPDI 208
DB 361 EVSSPPPTTEBVDHKKPLCLRLAGENGYGFLNALIRGLPGSFIKEVQKGP-----DL 414
QY 209 TALPDADRVTCGLCEGDGIRDI-----KMTVNRVPT 238
DB 415 AGLEBDEVIT--IEVNGVNVLDPEYERKVDRIQS 445

RESULT 6
US-09-461-474-6
Sequence 6, Application US/09461474
Patent No. 6278042
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Rafalecki, Antoni
TITLE OF INVENTION: Plant Metal Transporters
FILE REFERENCE: B1303 US NA
CURRENT APPLICATION NUMBER: US/09/461,474
CURRENT FILING DATE: 1999-12-14
EARLIER APPLICATION NUMBER: 60/112,562
```

EMBLER FILING DATE: 1998-12-16  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: Microsoft Office 97  
SEQ ID NO: 6  
LENGTH: 376  
TYPE: PPT  
ORGANISM: Triticum aestivum  
US-09-461-474-6

Query Match 5.3%, Score 88; DB 4; Length 376;  
Best Local Similarity 34.0%, Pred. No. 0.22;  
Matches 33; Conservative 15; Mismatches 35; Indels 14; Gaps 5;

QY 3 SURLSTHQSLLPSSLSQKTLISRFVNNPSPSPFS---VQPNRPPELAGETPR 52  
DB 12 SLLPFS-RHQ-----LRQRTL--QPHRTFRPKMSTAVISAEDALEPSSLQSLDORSIR 61

QY 60 IVVITSGKGVGKTTTANVGLSLARVGSVAIDAD 96  
DB 62 WIFV-GRKGVGRKTTTSSLAIQAKRPSVLLSTD 97

RESULT 7  
PCT-US95-17026-2  
Sequence 2, Application PC/TUS9517026

GENERAL INFORMATION:

APPLICANT: Zymogenetics, Inc.

APPLICANT: 1201 Eastlake Avenue East

APPLICANT: Seattle

APPLICANT: WA

APPLICANT: USA

APPLICANT: 98102

TITLE OF INVENTION: Bovine Factor XIII

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: ZymoGenetics, Inc.

STREET: 1201 Eastlake Avenue East

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98102

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/17026

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Parker, Gary E

REGISTRATION NUMBER: 31-648

REFERENCE/DOCKET NUMBER: 94-18PC

TELEPHONE: 206-442-6673

TELEFAX: 206-442-6678

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 732 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-17026-2

Query Match 5.3%, Score 88, DB 5, Length 732,  
Best Local Similarity 24.2%, Pred. No. 0.65,  
Matches 72; Conservative 38; Mismatches 94; Indels 94; Gaps 20;

QY 34 NSRRSPISVLOFNKPE---LAGETPRIVVITSGKGAVGKTTTANVGLSL----- 83  
DB 255 NPIKVSRSVGSAM-INKKDEGVIAAGSMNVVAY-----GVPPSAMTGSVDLLLEYKSSQK 308

QY 84 -ARYGSVAIDADLGLPNI-LLGLENPV-NTYCEVINQDRLD-----QA 129  
DB 309 FVRYGCCWVF---AGVFNFELRCLGIPAPVTVNYSADNDANLQDLFLEEDGNVSK 364  
QY 130 LVDPKPSNPFELCTSK---PSKLPWGFPGFALFMIYDALTPPEGSDPFIIDCPAGI 186  
DB 365 LTKDSVM-NHH--CWNAMMTRPDLPVGRG---WQV--VSTQENDGMYRCGPASV 415  
QY 187 -----DAGFITAITPANEAVLVTPDITALTARDARYTGLLEC----- 223  
DB 416 QAIKHGVCQFDAPFVFA--EVNSDLVY-----VFAKKDGTHTVEALDTTHIGKLVTK 469  
QY 224 ---GGIDIKMIVRVRTDMIK---GEDMSYLDVQEWLG---LSLGVFPEDSEV 270  
DB 470 TGSOSMKDI-----TDYKRFQEGQFEERLALETAMMYGAKKALNTEGVLSKSDV 519

RESULT 8  
US-08-597-236-5

Sequence 5, Application US/08597236

Patent No. 5733765

GENERAL INFORMATION:

APPLICANT: STINGELE, Francesca

APPLICANT: MOULET, Beat

TITLE OF INVENTION: LACTIC BACTERIA PRODUCING

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Penile & Edmonds

STREET: 1155 Avenue of the Americans

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/597,236

FILING DATE:

CLASSIFICATION: 426

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 95301669.9

FILING DATE: 20-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Panucci A., Allan

REGISTRATION NUMBER: 30256

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 249 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-597-236-5

Query Match 5.3%, Score 87; DB 1; Length 249;  
Best Local Similarity 20.8%; Pred. No. 0.15;  
Matches 48; Conservative 40; Mismatches 79; Indels 64; Gaps 10;

QY 41 IRSVLOFNKPE-----LAGETPRIVVITSGKGAVGKTTTANVGLSLARVGS 89  
DB 7 VSKVDFAKTEERYNARITNIGSQAGQMVIAISSVGESEKSMI SVLAIISFAVGAR 66  
QY 96 VVAIADAGLRLDLGLGLENRVITCEVYINQDCLDQALVADKRSNF-----EL 141  
DB 67 TLIDAE TEN -----SVLSGTFFSNEY...KGLSFLSGNADLNET 105  
QY 142 LC---IS-----PPSKLPWGFPGFALFMIYDALTPPEGSDPFIIDCPAG 185  
DB 106 ICQDTSILDVIAAGPVPNPFTSLQ---NDNPRHLMVARS---CYDYVILIDPPVG 157

QY 186 IDAGFITATPANEAVALTTPDITLRADRVGTGLLECDGIRDKIMTVNV 236  
DB 158 LVIDAVIIAHQADASLVTETGKIKPFTVTRAVEQLVESGSGFLGVNLNV 208

RESULT 9  
US-08-746-682A-5

Sequence 5, Application US/08746682A  
Patent No. 5786184  
GENERAL INFORMATION:  
APPLICANT: STINGELE, Francesca  
APPLICANT: MOLETT, Beat  
TITLE OF INVENTION: LACTIC BACTERIA PRODUCING  
TITLE OF INVENTION: EXOPOLYSACCHARIDES  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americans  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/746,682A  
FILING DATE: 14-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/597,236  
FILING DATE: 20-JUN-1995  
APPLICATION NUMBER: EP 95201669.9  
FILING DATE: 20-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fanucci A., Allan  
REGISTRATION NUMBER: 30256  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 249 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-746-682A-5

Query Match 5.3%; Score 87; DB 1; Length 249;  
Best Local Similarity 20.8%; Pred. No. 0.15;  
Matches 48; Conservative 40; Mismatches 79; Indels 64; Gaps 10;

QY 41 IRSVIFQFNKPE-----LAGETPRIVITSGKGVGKTTTANNGLSLARKGFS 89  
DB 7 VSKVDFPKFEEYVNAIPNTIQSGACMVIAISSVAGSGKPSMISVNAISFASVGLR 66  
QY 90 VVAIDADGLRNI.DLLGLRNVVYTCVINGDCPLDQALVPRKMSNF-----EL 141  
DB 67 TLLIDAE--TRN-----SVLSGTFKSNRPY--KGLSNFLSGNADLNFT 105  
QY 142 LC---IS-----KRSKLPKMGFGSKALEMLVDALKTRPESSPFIITDC-PAG 185  
DB 106 ICQTDISGLDVASGPVPNPSTSLQ---NDNFRHLMVARS---CYDYVIIDTPPVG 157  
QY 186 IDAGFITATPANEAVALTTPDITLRADRVGTGLLECDGIRDKIMTVNV 236  
DB 158 LVIDAVIIAHQADASLVTETGKIKRFRVTYKAVEQLVESGSGFLGVNLNV 208

RESULT 10  
US-09-320-878-18  
Sequence 18, Application US/09320878A  
Patent No. 6117659  
GENERAL INFORMATION:

APPLICANT: ASHLEY, Gary  
APPLICANT: BETLACH, Melanie C.  
APPLICANT: BETLACH, Mary C.  
APPLICANT: MCDANIEL, Robert  
APPLICANT: TANG, Li

## TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE

FILE REFERENCE: 30062202120  
CURRENT APPLICATION NUMBER: US/09/320,878A  
CURRENT FILING DATE: 1999-05-27  
EARLIER APPLICATION NUMBER: CIP OF 09/141,908  
EARLIER FILING DATE: 1998-08-28  
EARLIER APPLICATION NUMBER: CIP OF 09/073,538  
EARLIER FILING DATE: 1998-05-06  
EARLIER APPLICATION NUMBER: CIP OF 08/846,247  
EARLIER FILING DATE: 1997-04-30  
EARLIER APPLICATION NUMBER: 60/119,139  
EARLIER FILING DATE: 1999-02-08  
EARLIER APPLICATION NUMBER: 60/100,980  
EARLIER FILING DATE: 1998-09-22  
EARLIER APPLICATION NUMBER: 60/087,080  
EARLIER FILING DATE: 1998-05-28  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 18  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-09-320-878-18

Query Match 5.2%; Score 86.5; DB 3; Length 416;  
Best Local Similarity 22.2%; Pred. No. 0.39;  
Matches 42; Conservative 31; Mismatches 73; Indels 43; Gaps 7;

QY 125 PLDQALVDRKMSNFFELCTSKPPSKLPWFGSKALEMLVDALKTRPESSPFIITDCPA 184  
DB 102 RLKRVAREFTMRVELL---RPRVQ-----EIVDGLVDMLAAPDGRADLM----- 145  
QY 185 CIDAGFITATPANEAVALTTPDITLR-----DABRVGTGLLECDGIRDKIMTV 233  
DB 146 -ESLAMPPIVISELGLVPEPRAAFRWTDGFVPPDDPAQAQTMAEMSGY--LSRLI 202  
QY 234 NRVPTDMIKGEDMSV-----LDVQEWLGLSLGVIPDESEVTRTNRGFFLV 282  
DB 203 DSKRGQ--DGEDDLALVRSDSDGSLTSEELGKAHILVVAHETTIVNLIANGMYALL 260  
QY 283 NKPPTLAGL 291  
DB 261 SHPDQLAAL 269

RESULT 11  
US-09-105-537-39  
Sequence 39, Application US/09105537A  
Patent No. 6265202  
GENERAL INFORMATION:

APPLICANT: Sherman, D.H.  
APPLICANT: Liu, H.  
APPLICANT: Xue, Y.

APPLICANT: Zhao, L.  
TITLE OF INVENTION: DNA encoding methymycin and pikromycin

FILE REFERENCE: 600.438US1  
CURRENT APPLICATION NUMBER: US/09/105,537A  
CURRENT FILING DATE: 1998-06-26

NUMBER OF SEQ ID NOS: 43  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 39  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-09-105-537-39

Query Match 5.2%; Score 86.5; DB 4; Length 416;  
Best Local Similarity 22.2%; Pred. No. 0.39;



Patent No. 5233423  
 APPLICANT: FRANCHINI, GENOVEFFA, MONG-STAL, FLOSSIE,  
 GALLO, ROBERT  
 TITLE OF INVENTION: CHARACTERIZATION OF REPLICATION COMPETENT  
 HUMAN IMMUNODEFICIENCY TYPE 2 PROVIRAL CLONE HIV-2 SBI/15Y  
 NUMBER OF SEQUENCES: 4  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/331,212  
 FILING DATE: 03-31-1989  
 SEQ ID NO: 2  
 LENGTH: 3077  
 5233423-2

Query Match 5.0%; Score 83.5; DB 6; Length 3077;  
 Best Local Similarity 21.5%; Pred. No. 21;  
 Matches 61; Conservative 48; Mismatches 122; Indels 53; Gaps 15;

QY 57 TTRIVITSGKGVGKTTTANVGLSLARYGFSVAIDADLGLNLLGLEPNVNTVC 116  
 DB 2037 TPLCVTMNOMASTESAVATTSRSGPDMINDPCLQANCGLEEDM---VECOFMMTG 2093  
 QY 117 VEYINGDCRLDAIVDPKWSNFELLG---ISKPEKLPKMGFGKAL-----EWLVD 165  
 DB 2094 LE-----LDKKKOYSEWYSKDVVCESDNSTDKRCYNNHCNTSVITESCDGHY--D 2144  
 QY 166 ALKTPRESGPDIIIDCPAGIDAGFITAITANEAVLVTPDITALADADRVGELLECDG 225  
 DB 2145 AMRFYCAPGVLKRCNDNTYSGF---EPVCSKVAVST---CTRMETOPSTWL---G 2194  
 QY 226 IPD1KMIIVNRPVD---MIKEDMSVLVDQEMIGLGLVPEDESEVIRST--NRG 277  
 DB 2195 F-----NGTPAFNPPTYVYHGRPNPTTISLNKYVNLITLCPPFENHTVVPITLMSGR 2247  
 QY 278 F--PLVKNPPTLAGLAFQAAKRLVEDOSKMAVVEEPEPKRG 319  
 DB 2248 FHSQYITNPPQAWCP-KGEPEAMQF VKQTLV-KHPRYKG 2288

## RESULT 15

US-08-416-603-4  
 Sequence 4, Application US/08416603  
 Patent No. 5866780  
 GENERAL INFORMATION:  
 APPLICANT: Law, Marcus  
 APPLICANT: Hebara, Ledare  
 APPLICANT: Reddick, Bradford B.  
 TITLE OF INVENTION: Maize Chlorotic Dwarf Virus Genome and  
 NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Saliwanichik & Saliwanichik  
 STREET: 2421 N.W. 41st Street, Suite A-1  
 CITY: Gainesville  
 STATE: FL  
 COUNTRY: USA  
 ZIP: 32606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30B  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/416,603  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lloyd, Jeffrey  
 REGISTRATION NUMBER: 35,589  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 904-375-8100  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3457 amino acids

TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-416-603-4

Query Match 5.0%; Score 83.5; DB 2; Length 3457;  
 Best Local Similarity 25.1%; Pred. No. 26;  
 Matches 53; Conservative 30; Mismatches 87; Indels 41; Gaps 14;

QY 18 SLSQKTLISSPRFVNPNRRSPRISVL-QFNKPE-LAGETPRIVITSGKGVGKTTT 75  
 DB 2888 TSAATKTSI-----RKSPIHGLVGEITPTEPSLIHAHDPPL-----PDXIGKMP 2932  
 QY 76 TANVGLSLARYGFSVAIDADLGLNLLGLEPNVNTVC---EV-INGDCRLD 127  
 DB 2933 VLEASM---KYGSRITFPFVDQILEVEDHLSKMLANCENSKNRQVNNLEIGING--ID 2986  
 QY 128 CALVPDPRWSNFELLGTSK-PPSK-LPMFGGYAL-EWLVDALTPREGSPDIIIDCPAG 185  
 DB 2987 GS---DYWQOIEMDTSSGWPYAKRKPVGAGK--KWLFEQDGTYPGKPRYVFGD--AG 3038  
 QY 186 IDAGFITAITPANFAVLVTPDITALADAD 216  
 DB 3039 LIESYNSMLGEAKGKGISPTVTITLCAKADERR 3069

Search completed: April 16, 2003, 09 03.02  
 Job time: 21 secs





GenCore version 5.1.4, p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 08:59:57 ; Search time 39 Seconds  
(without alignments)  
1113.839 Million cell updates/sec

Title: US-10-067-989-1  
Perfect score: 1654  
Sequence: 1 MASLPIRSTNMQSLIPSSL KAVMEEPYRPGFPFGG 326

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: A.GeneSeq.101002:.\*  
2: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.\*  
3: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.\*  
4: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.\*  
5: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.\*  
6: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.\*  
7: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.\*  
8: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.\*  
17: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.\*  
18: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.\*  
19: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.\*  
20: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.\*  
21: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.\*  
22: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.\*  
23: /SID22/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1654	100.0	326	21	AAAG6531
2	1654	100.0	326	21	AAAG7959
3	1654	100.0	343	21	AAAG6530
4	1648	99.6	326	21	AAAG1292
5	1648	99.6	344	21	AAAG1292
6	1175.5	71.1	295	23	AAAG1292
7	888	53.7	174	21	AAAG1292
8	888	53.7	174	21	AAAG6532
9	545.5	33.0	266	23	AAAB9459
10	365	22.1	260	22	AAAB9685

11	327.5	19.8	245	22	AAAB96374
12	252.5	15.3	258	22	AAAB96719
13	224	13.5	294	19	AAAY10300
14	224	13.5	297	18	AAAB20738
15	209	12.6	425	18	AAAB21013
16	206	12.5	388	21	AAAB1657
17	202	12.2	300	22	AAAB96246
18	198	12.0	412	18	AAAB20443
19	197.5	11.9	279	22	AAAG93171
20	196.5	11.9	205	22	AAAB96618
21	190	11.5	390	22	AAAG81141
22	183	11.1	313	21	AAAG11242
23	183	11.1	313	21	AAAG43414
24	178	10.8	289	21	AAAG43419
25	176	10.6	288	21	AAAB9622
26	173.5	10.5	375	22	AAAG90990
27	171.5	10.4	271	22	AAAB92570
28	165	10.0	115	21	AAAB40393
29	165	10.0	115	23	ABP03821
30	164.5	9.9	478	22	AAAG90129
31	163	9.9	293	22	ABAB58650
32	163	9.9	339	22	AAU39817
33	161	9.7	260	22	AAAB71330
34	161	9.7	400	21	AAAG22284
35	161	9.7	400	21	AAAG39877
36	161	9.7	532	21	AAAG22283
37	161	9.7	532	21	AAAG39876
38	161	9.7	547	21	AAAG39875
39	161	9.7	548	21	AAAG22282
40	159.5	9.6	257	21	AAAT75106
41	158.5	9.6	257	21	AAAT75107
42	157.5	9.5	286	22	AAAB96245
43	154.5	9.3	350	21	AAAG41434
44	152.5	9.2	296	21	AAAG41435
45	152	9.2	342	23	AAAB47911

## ALIGNMENTS

RESULT 1  
ID AAAG6531  
AAAG6531 standard; Protein; 326 AA.  
XX  
AC AAAG6531;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58550.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
XX  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135355.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137722.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139762.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139889.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144088.  
PR 16-JUL-1999; 99US-0144088.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159325.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.

```

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 100.0%; Score 1654; DB 21; Length 326;  
 Best Local Similarity 100.0%; Pred. No. 7, 3e-174;  
 Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MASLRFSTNHQSLLPSSLSQKTLISSPFVNNPSPSPISVLAQFNKPELAGETPTI 60
DB 1 MASLRFSTNHQSLLPSSLSQKTLISSPFVNNPSPSPISVLAQFNKPELAGETPTI 60
QY 61 VVITSGKGVGKTTTANVGLSLARVGFVSVAIDADLGRNLDLLGLENRVNYTCVEVI 120
DB 61 VVITSGKGVGKTTTANVGLSLARVGFVSVAIDADLGRNLDLLGLENRVNYTCVEVI 120
QY 121 NGDCRLDQALVRDPRWNSFELLCTISKPRSKLPMGFGKALEMVLVDALKTRPESGDPFII 180
DB 121 NGDCRLDQALVRDPRWNSFELLCTISKPRSKLPMGFGKALEMVLVDALKTRPESGDPFII 180
QY 181 DCPAGIDAGFTTATTPANEAVALVTTPTDITALRDADRVYTGLECDGIRDIKMIIVNRRTDM 240
DB 181 DCPAGIDAGFTTATTPANEAVALVTTPTDITALRDADRVYTGLECDGIRDIKMIIVNRRTDM 240
QY 241 IKGEDMMSVLDVQEMGLSLGVIPEDSEVIRSTNRGFPVLVUNKPPTIAGLAFEDQAAARL 300
DB 241 IKGEDMMSVLDVQEMGLSLGVIPEDSEVIRSTNRGFPVLVUNKPPTIAGLAFEDQAAARL 300
QY 301 VEQDSMKAVWVEEPEKRGFFSFFCG 326
DB 301 VEQDSMKAVWVEEPEKRGFFSFFCG 326

```

## RESULT 2

AAG77959  
 ID AAG77959 standard; Protein; 326 AA.

AAAG77959;

01-FEB-2002 (first entry)

Arabidopsis thaliana AtMind1 protein.

Mind; AtMind1; chloroplast; plastid; photosynthesis.

Arabidopsis thaliana.

WO200181601-A2.

01-NOV-2001.

18-APR-2001; 2001WO US:2634.

XX

PR 19-APR-2000; 2000US-0553431.

XX (UYNE-) UNIV NEVADA & COMMUNITY COLLEGE SYSTEM.

XX Osteryoung KW;

XX WPI: 2002-041414/05.

XX N-PSDB; AAH77248.

XX

XX

PS Example; Fig 1; 35pp; English.

The sequence represents A. thaliana AtMind1. The invention relates to a DNA sequence isolated from its native genome comprising a plant Mind gene that functions in regulating plastid division. When ectopically expressed, the gene alters the number, shape and/or size of chloroplasts and other types of plastids present in plant cells. A genetic construct comprising the gene of the invention in sense or antisense orientation is useful for altering the size, shape and/or number of plastids, in particular chloroplasts in plant cells. Alteration in the plastid size, shape and number improves productivity or increases vigour due to enhanced photosynthetic capacity and allows enhanced production of commercially important compounds that accumulate naturally or as a result of genetic engineering.

XX Sequence 326 AA;

Query Match 100.0%; Score 1654; DB 23; Length 326;  
 Best Local Similarity 100.0%; Pred. No. 7, 3e-174;  
 Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MASLRFSTNHQSLLPSSLSQKTLISSPFVNNPSPSPISVLAQFNKPELAGETPTI 60
DB 1 MASLRFSTNHQSLLPSSLSQKTLISSPFVNNPSPSPISVLAQFNKPELAGETPTI 60
QY 61 VVITSGKGVGKTTTANVGLSLARVGFVSVAIDADLGRNLDLLGLENRVNYTCVEVI 120
DB 61 VVITSGKGVGKTTTANVGLSLARVGFVSVAIDADLGRNLDLLGLENRVNYTCVEVI 120
QY 121 NGDCRLDQALVRDPRWNSFELLCTISKPRSKLPMGFGKALEMVLVDALKTRPESGDPFII 180
DB 121 NGDCRLDQALVRDPRWNSFELLCTISKPRSKLPMGFGKALEMVLVDALKTRPESGDPFII 180
QY 181 DCPAGIDAGFTTATTPANEAVALVTTPTDITALRDADRVYTGLECDGIRDIKMIIVNRRTDM 240
DB 181 DCPAGIDAGFTTATTPANEAVALVTTPTDITALRDADRVYTGLECDGIRDIKMIIVNRRTDM 240
QY 241 IKGEDMMSVLDVQEMGLSLGVIPEDSEVIRSTNRGFPVLVUNKPPTIAGLAFEDQAAARL 300
DB 241 IKGEDMMSVLDVQEMGLSLGVIPEDSEVIRSTNRGFPVLVUNKPPTIAGLAFEDQAAARL 300
QY 301 VEQDSMKAVWVEEPEKRGFFSFFCG 326
DB 301 VEQDSMKAVWVEEPEKRGFFSFFCG 326

```

## RESULT 3

AAG46530  
 ID AAG46530 standard; Protein; 343 AA.

AAAG46530;

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 58549.

Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

XX

OS Arabidopsis thaliana.  
XX  
XX EPI033405 A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301432.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130444.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.

FR 30-JUN-1999; 99US-0141287.  
FR 01-JUL-1999; 99US-0141642.  
FR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
FR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
FR 09-JUL-1999; 99US-0142820.  
FR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
FR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
FR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
FR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
FR 03-AUG-1999; 99US-0147038.  
FR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147392.  
PR 05-AUG-1999; 99US-0147393.  
PR 06-AUG-1999; 99US-0147393.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
FR 20-AUG-1999; 99US-0149929.  
FR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.





|||||  
Db 181 DCPAGIDAGFIATIPANNAVLTTPDITALPDADRVIGLLECDIIRIKMIVNRVRTDM 240  
OY 241 IKGEDMMSVLDVQEMHGLSLLEVIPEDESVIRSTRNGPPLVINKPPTLAGAFEQAAARL 300  
Db 241 IKGEDMMSVLDVQEMHGLSLLEVIPEDESVIRSTRNGPPLVINKPPTLAGAFEQAAARL 300  
OY 301 VEQDSMKAVMVEEPPKRGFSFPGG 326  
Db 301 VEQDSMKAVMVEEPPKRGFSFPGG 326  
RESULT 5  
AAG12921  
ID AAG12921 standard, Protein: 344 AA.  
XX  
AC AAG12921;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 12221.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hydridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139482.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.



```

PR 09-AUG-1999; 99US-0147493
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 99.6%; Score 1648; DB 21; Length 344;
Best Local Similarity 99.7%; Pred. No. 3.6e-173;
Matches 325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MASLPSTNHSQILLPSSLSQKTLISSPPFVNNPSPSPISRSV. QPNKKPELAGTPTI 60
DB 19 MASLPSTNHSQILLPSSLSQKTLISSPPFVNNPSPSPISRSV. LQFNKKPELAGTPTI 78
QY 61 VVITSGKGVGKTTTANVGLSLAPYGVVAIDAN/GI/PNDLLLLGHPNPNYTCVEVI 120
DB 79 VVITSGKGVGKTTTANVGLSLAPYGVVAIDAN/GI/PNDLLLLGHPNPNYTCVEVI 138
QY 121 NGDCRLDQALVDRKMSNFELLCISKRSKLPWFGSKALEMLVDALKTPPGSPDF111 180
DB 139 NGDCRLDQALVDRKMSNFELLCISKRSKLPWFGSKALEMLVDALKTPPGSPDF111 198
QY 191 DCPAGIDAGFTTATTPANPAVLVTPPTITLBPADPVTGLLETCGIRPDKMIVNRPPTM 240
DB 199 DCPAGIDAGFTTATTPANPAVLVTPPTITLBPADPVTGLLETCGIRDKMIVNRPPTM 258
QY 241 IKGEDPMSVLDVQEMLGSLGVIPEDSEVIPTSTNGFPLVINKPPTLAGLAFEQAMPL 300
DB 259 IKGEDPMSVLDVQEMLGSLGVIPEDSEVIPTSTNGFPLVINKPPTLAGLAFEQAMPL 318
QY 301 VEQDSMKAVWEPEPKKRGFPSPFPG 326
DB 319 VEQDSMKAVWEPEPKKRGFPSPFPG 344

RESULT 6
AAG77960
ID AAG77960 standard; Protein; 295 AA.
XX
AC AAG77960;
XX
DT 01-FEB-2002 (first entry)
XX
DE Tagetes erecta Temind protein.
XX
KW African marigold; Mind; Temind; chloroplast; plastid;
XX photosynthesis.
XX
OS Tagetes erecta.
XX
MO200181601-A2.
XX
PD 01-NOV-2001.
XX
PF 18-APR-2001; 2001WO-US:2634.
XX
PR 19-APR-2000; 2000US-0553431.
XX
PR 19-APR-2000; 2000US-0553431.
XX
PA (UNNE-) UNIV NEVADA & COMMUNITY COLLEGE SYSTEM.
XX
PI Osteeryoung KW;
XX
DR WPI; 2002-041414/05.
XX
DR N-PSDB; AAH77249.
XX
PT Novel DNA sequence involved in regulating plastid division in plants
XX useful for obtaining transgenic plants with novel phenotypes,
XX characterised by alterations in plastid shape, size and/or number
XX
PS Example; Page 34-35; 35pp; English.
XX
CC The sequence represents T. erecta Temind protein. The invention relates
CC to a DNA sequence isolated from its native genome comprising a plant
CC Mind gene that functions in regulating plastid division. When ectopically
CC expressed, the gene alters the number, shape and/or size of chloroplasts
CC and other types of plastids present in plant cells. A genetic construct
CC comprising the gene of the invention in sense or antisense orientation is
CC useful for altering the size, shape and/or number of plastids, in
CC particular chloroplasts in plant cells. Alteration in the plastid size,
CC shape and number improves productivity or increases vigour due to
CC enhanced photosynthetic capacity and allows enhanced production of

```

CC commercially important compounds that accumulate naturally or as a result  
of genetic engineering.  
XX  
SQ Sequence 295 AA;  
Query Match 71.1%; Score 1175.5, DB 23, Length 295;  
Best Local Similarity 72.0%; Pred. No. 56-121;  
Matches 247; Conservative 15; Mismatches 16; Indels 65; Gaps 7;  
QY 1 MASLRFTSNHOSLLPSSLSQKTLISSPRPVN-NPSRRS-----PIRSVLQ 46  
DB 1 MTSLR-----LTERSLVCSSTFPFNFHLKTKPRPKPKPRKPIRSVLQ 47  
QY 47 FNRKPELAGETPRIVITSGKGVGKTTTAVNG:SIAPYGVSAVADLGLPNDLL 106  
DB 48 YNKKPELAGETPR-----VVAIDADVGLRNDLL 77  
QY 107 GLENRVNVCVEVINGDCRLDQALVPDKPWSNPELCTISKPSKTPMGSGKALFEMLVDA 166  
DB 78 GLENRVNVTVEVLNGDCRLDQALVPDKPWSNPELCTISKPSKTPMGSGKALFEMLVDA 136  
QY 167 LKTRPESGSDFTIIDCPAGIDAGFTAITPANEAVLVTPDITLADPVTGLLECDGI 226  
DB 137 LKDRQSGCPDFTIIDCPAGIDAGFTAITPANEAVLVTPDITLADPVTGLLECDGI 196  
QY 227 PDIKMTVNPPTMTMKGEMMSVLDVQEMLGSLGVTPEDS---PVIPSTNPGEPIVYN 283  
DB 197 RDKMTVNPPTDITIRGEMMSVLDVQEMLGSLGVTPEDS---SDIRGFEVIRKSTNKGPELVIN 252  
QY 284 KPPTLAGIAFEQAAMPILVQNSKAVMVEEPEKSGFSPFQ3 326  
DB 253 KPPTLAGIAFEQAAMPILVQNSKAVMVEEPEKSGFSPFQ3 295  
RESULT 7  
AA012923  
ID AA012923 standard; Protein; 174 AA.  
XX  
AC AA012923;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 12223.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128334.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR

PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0140039.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143154.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR

```

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145182.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145216.
PR 23-JUL-1999; 99US-0145274.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145911.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146385.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153768.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158929.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

```

```

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 53.7%; Score 888; DB 21; Length 174;  
 Best Local Similarity 100.0%; Pred. No. 1,3e-89; Mismatches 0; Indels 0; Gaps 0;  
 Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 153 MGFGKALEMLVDALKTRPESSPFIIDCPAGIDAGFIATTPANEAVLVTTFD;TALR 212
DB 1 MGFGKALEMLVDALKTRPESSPFIIDCPAGIDAGFIATTPANEAVLVTTFD;TALR 60

```

```

QY 213 DADRTGLEECGIPDTIMVNPPTTMYKGEEMSVLNQEMLSLSLVTPENSEVIR 272
DB 61 DADRTGLEECGIPDTIMVNPPTTMYKGEEMSVLNQEMLSLSLVTPENSEVIR 120

```

```

QY 273 SNPGFFLVNKKPPTLAGLAFCGAAPLVEDNSKAVMVEEKKGFPSFPQ 326
DB 121 SNPGFFLVNKKPPTLAGLAFCGAAPLVEDNSKAVMVEEKKGFPSFPQ 174

```

#### RESULT 8

ID AAG46532 standard, Protein; 174 AA.

XX AAG46532;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 58551.

XX Protein identification; signal transduction pathway; metabolic pathway;

KM hydridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP 2000.

PF 25-FEB 2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.



PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 53.7%; Score 888; DB 21; Length 174;

Best Local Similarity 100.0%; Pred. No. 1,3e-89; Mismatches 0; Indels 0; Gaps 0;

Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 153 MGFSGKALEMVLVDALKTREPSDFIICDPAGIDAGFITAITPANEAVLTPDITAIR 212  
 DB 1 MGFSGKALEMVLVDALKTREPSDFIICDPAGIDAGFITAITPANEAVLTPDITAIR 60  
 OY 213 DADRITGLLECGITPDITMYNVPVPTMTFGEEMSVLDVDEMGLSLVYIPEDSEVIR 272  
 DB 61 DADRITGLLECGITPDITMYNVPVPTMTFGEEMSVLDVDEMGLSLVYIPEDSEVIR 120  
 OY 273 STNRGFPVLNKPPLIAGLAFEOAAMPLEVDGSMKAWVEEPPYPPFPSPFG 326  
 DB 121 STNRGFPVLNKPPLIAGLAFEOAAMPLEVDGSMKAWVEEPPYPPFPSPFG 174

RESULT 9  
 ABR49459  
 ID ABR49459 standard; Protein; 266 AA

XX ABR49459;

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #2163.

XX Antihacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
 KW vitamin B12; bacterial infection; disease.

XX Listeria monocytogenes.

OS WO20017735-A2.

PN 18-OCT-2001.

XX 11-APR-2001; 2001WO-FR01118.

PF 11-APR-2000; 2000FP-0004629

XX (INSP) INST PASTEUR.

PI Buchrisser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P,  
 PI Duesurget O, Chetouani F, Nedjati H, Glaser P, Kunst F, Cossart P,  
 PI Daniels J, Goebel W, Kretz J, Kuhn M, Ng E, Vazquez-Boland JA,  
 PI Dominguez-Bernal G, Garrido-Garcia P, Tietz-Warltner A, Amend A,  
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L,  
 PI Perez-Diaz J, Baguero F, Garcia Del Portillo F, Gomez-Lopez N,  
 PI Maduenio E, De Fabios B, Wendland J, Kaerst U, Entian K, Hauf J,  
 PI Rose M, Voss H,

DR WPI, 2002-010914/01.

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment

PT and prevention of Listeria and related bacterial infections, and  
 PT related polypeptides  
 XX Claim 6; SEQ ID No 2164; 192pp; French.

CC The present invention relates to the genome sequence of Listeria  
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of  
 CC it are useful for selecting probes and primers for detecting genes in L.  
 CC monocytogenes and related organisms, and for studying genetic  
 CC polymorphisms and other genomes. The present sequence is a protein  
 CC encoded by the genome sequence of the present invention. Proteins  
 CC expressed from the genome sequence are useful for raising specific  
 CC antibodies, identification of L. monocytogenes and related organisms,  
 CC for biosynthesis and biodegradation, especially biosynthesis of vitamin  
 CC B12. The genome sequence and proteins encoded by it are also useful for  
 CC selecting compounds that regulate gene expression and cell replication  
 CC and modulate L. monocytogenes-related diseases. In addition, the genome  
 CC sequence and proteins encoded by it are useful in pharmaceutical and  
 CC vaccine compositions for the treatment or prevention of infections by L.  
 CC monocytogenes and related organisms.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pot\_sequences.

SQ Sequence 266 AA;

Query Match 33.0%; Score 545.5; DB 23; Length 266;  
 Best Local Similarity 43.9%; Pred. No. 1,3e-52; Mismatches 115; Conservative 59; Mismatches 79; Indels 9; Gaps 5;

OY 61 VITSGKGGVKTTPITAVGLSLARVGSVAIDALGRNLDLGLBNVYTCVEVI 120  
 DB 5 VITSGKGGVKTTPITAVGLSLARVGSVAIDALGRNLDLGLBNVYTCVEVI 64

OY 121 NGDCRPLDQALVDRKPSNPELCTSPRSKPLPMGFYKALEMVLVDALKTREPSDFI 180

DB 65 EGRCYKHQAMIKRPFDDLTFLPAQTTDKNAVSGEQWDLINQL--PDD--YDFILI 119

OY 181 DCPAGIAGFITAITPANEAVLTPDITPAPASVLTLETPGIPITMYVNPVPTDM 240

DB 120 DCPAGIETGYKNAVAGADKRAIVVTPPISAVPDADPIIGULEKEDIPEPKLIINPIPTQM 179

OY 241 IKGEDMMSVLVDDEMGLSLVYIPEDSEVIRSTNGFPLVINKPPTLGLAFEOAAML 300

DB 180 MNGDVMDIDITTHLSIELGLIIDDDEVIRSSNGDPVAM-LPNRMSQCYPNIAPIRI 238

OY 301 VEQDSMKAWVEEPPYPPFPSPFG 322

DB 239 LGE--SIPLMSIETKXAGFFA 257

RESULT 10

AB96685  
 ID AAB96685 standard; Protein; 260 AA.

XX AAB96685;

DT 29-OCT-2001 (first entry)

DE Putative P. abyssus ATPase involved in chromosome partitioning #6.

XX Hyperthermophilic archaeon; hyperthermophilic protein.

OS Pyrococcus abyssi.

XX FR2792651-A1.

XX 27-OCT 2000.

XX 21-APR-1999; 99FR-0005034.

XX 21-APR-1999; 99FR-0005034.

PA (CNRS ) CNRS CENT NAT RECH SCI.  
 PA (IFRE-) IFREMER INST FR RECH EXPL MER.  
 PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;  
 PI Querellou J, Weissenbach J, Saurin W, Hellig R;  
 DR WPI: 2001-126236/14.  
 XX  
 PT New nucleotide sequences isolated from *Pyrococcus abyssi* encode  
 PT proteins useful in industry -  
 XX  
 PS Claim 7; Pages 1443-1444; 1657pp; French.  
 CC  
 CC The present invention relates to the genomic sequence of *Pyrococcus*  
 CC *abyssi* (see AAF66431 and AAH41223-7) and *P. abyssi* proteins. *P. abyssi* is  
 CC a hyperthermophilic archaeon, which is isolated from deep-sea  
 CC hydrothermal vents. The present sequence is one such *P. abyssi* protein.  
 CC The proteins of the present invention have various potential industrial  
 CC uses, since the proteins are stable at very high temperatures, some up to  
 CC 110 degrees centigrade.  
 CC Note: This patent is in the same patent family as WO200065062, which  
 CC contains additional sequences as shown in AAB99132-AAB99143,  
 CC AAH75903-AAH75920 and AAG66436.  
 CC  
 SQ Sequence 260 AA;

Query Match 22.1%; Score 366; DB 22; Length 260;  
 Best Local Similarity 32.9%; Pred. No 1.3e-31;  
 Matches 94; Conservative 54; Mismatches 88; Indels 50; Gaps 7;

OY 59 RIVVITSGKGGVGTATTANVGI.SLARVGSVVAIDAPLGLNLDLGLERNVNYCWE 118  
 DB 4 RIVVITSGKGGVGTATTANVGI.SLARVGSVVAIDAPLGLNLDLGLERNVNYCWE 118  
 OY 119 VINQCPDLQALVPDKPWSNELLCTSKPSKLPWG- - -GHALEMLVDAKYTPPEG- 173  
 DB 63 VLAGADLKDIAIYEG- - - - -PAGVVIIPGSLSEKIKKAAEAERLPDL 104  
 OY 174 - - - - -SPDFIIDCPAGIDAGFITAITPANEAVLVTPDTALRDADRVTLGLECCGIR 227  
 DB 105 IREISQMDPILIDAPAGLELTSTALIGKELIVTNPEIAITDSLKTKLVAEKLGL 164  
 OY 228 DTKMIVNEPPTDMKSGDMSVLQCEMLGLSLGLVLPEDSEVTPSTNPPFVLNKPPT 287  
 DB 165 PLGALINVTSEKTE- - - - -LSPEEIEALEVPVLTGTVEDPEVBAASAVGLVVKPPTS 220  
 OY 288 IAGLAFEGQA- - - - -WPIVQDSMAVAVVEEPPYRGSPFFG 326  
 DB 221 PAIAVYFIAKLACIWKPRPEPES- - - - -PVYPIFPALFG 257

RESULT 11  
 AAB96374  
 ID AAB96374 standard; Protein; 245 AA.  
 AC AAB96374;  
 XX  
 DT 29-OCT-2001 (first entry)  
 DE Putative P. *abyssi* ATPase involved in chromosome partitioning #3.  
 XX  
 KW Hyperthermophilic archaeon; hyperthermophilic protein.  
 XX  
 OS *Pyrococcus abyssi*.  
 XX  
 PN FR2792651-A1.  
 PD 27-OCT-2000.  
 XX  
 PF 21-APR-1999; 99FR-0005034.  
 XX  
 PR 21-APR-1999; 99FR-0005034.  
 XX

PA (CNRS ) CNRS CENT NAT RECH SCI.  
 PA (IFRE-) IFREMER INST FR RECH EXPL MER  
 XX  
 PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;  
 PI Querellou J, Weissenbach J, Saurin W, Hellig R;  
 DR WPI: 2001-126236/14.  
 XX  
 PT New nucleotide sequences isolated from *Pyrococcus abyssi* encode  
 PT proteins useful in industry -  
 XX  
 PS Claim 7; Pages 1058-1059; 1657pp; French.  
 CC  
 CC The present invention relates to the genomic sequence of *Pyrococcus*  
 CC *abyssi* (see AAF66431 and AAH41223-7) and *P. abyssi* proteins. *P. abyssi* is  
 CC a hyperthermophilic archaeon, which is isolated from deep-sea  
 CC hydrothermal vents. The present sequence is one such *P. abyssi* protein.  
 CC The proteins of the present invention have various potential industrial  
 CC uses, since the proteins are stable at very high temperatures, some up to  
 CC 110 degrees centigrade.  
 CC Note: This patent is in the same patent family as WO200065062, which  
 CC contains additional sequences as shown in AAB99132-AAB99143,  
 CC AAH75903-AAH75920 and AAG66436.  
 CC  
 SQ Sequence 245 AA;

Query Match 19.8%; Score 327.5; DB 22; Length 245;  
 Best Local Similarity 34.6%; Pred. No. 2.1e-27;  
 Matches 90; Conservative 49; Mismatches 90; Indels 31; Gaps 8;

OY 59 RIVVITSGKGGVGTATTANVGI.SLARVGSVVAIDAPLGLNLDLGLERNVNYCWE 118  
 DB 3 RIVVITSGKGGVGTATTANVGI.SLARVGSVVAIDAPLGLNLDLGLERNVNYCWE 118  
 OY 119 VINQCPDLQALVPDKPWSN- - - - -FELLCTSKPSKLPWGFGKALEMLVDAKYT 163  
 DB 62 VLAGADLKDIAIYV- - - - -TQFDNVVYLPAGVDEHVMKADPR-KLPE- - - - -VTKSLK- 109  
 OY 170 RPEGSPDITIDCPAGIDAGFITAITPANEAVLVTPDTALRDADRVTLGLECCGIRDI 229  
 DB 110 - - - - -GDYDFILDCPAGIQLDAMSAMLSGEALIVTNPEISCLDTMKVGIIVRKAGLAIL 166  
 OY 230 KIVNRY-PTCMIGEMPMGLVQCEMLGLSLGLVLPEDSEVTPSTNPPFVLNKPPTL 288  
 DB 167 GFILNRKRIENDLPFA- - - - -AQDMVDFLAVLPEDPAIREGLTIGIIVVKYRPESE 221  
 OY 289 AGLAFEGQAAMRLVQDSMA 308  
 DB 222 GAKAFVLAQVEKLAICIK 241

RESULT 12  
 AAB96719  
 ID AAB96719 standard; Protein; 258 AA.  
 AC AAB96719;  
 XX  
 DT 29-OCT-2001 (first entry)  
 DE Putative P. *abyssi* ATPase involved in chromosome partitioning #7.  
 XX  
 KW Hyperthermophilic archaeon; hyperthermophilic protein.  
 XX  
 OS *Pyrococcus abyssi*.  
 XX  
 PN FR2792651-A1.  
 PD 27-OCT-2000.  
 XX  
 PF 21-APR-1999; 99FR-0005034.  
 XX  
 PR 21-APR-1999; 99FR-0005034.  
 XX

PA (CNRS ) CNRS CENT NAT RECH SCI.  
 PA (IPRE-) IPRMER INST FR RECH EXPL MER.  
 XX  
 XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;  
 PI Querellou J, Weissenbach J, Saurin W, Hellig R;  
 XX WPI: 2001:126236/14.  
 DR  
 XX  
 PT New nucleotide sequences isolated from *Pyrococcus abyssi* encode  
 PT proteins useful in industry -  
 XX  
 PS Claim 7, Pages 1481-1482, 1657pp, French.  
 XX  
 CC The present invention relates to the genomic sequence of *Pyrococcus*  
 CC *abyssi* (see AAF6431 and AAF4123.7) and *P. abyssi* proteins. *P. abyssi* is  
 CC a hyperthermophilic archaeon, which is isolated from deep-sea  
 CC hydrothermal vents. The present sequence is one such *P. abyssi* protein.  
 CC The proteins of the present invention have various potential industrial  
 CC uses, since the proteins are stable at very high temperatures, some up to  
 CC 110 degrees centigrade.  
 CC Note. This patent is in the same patent family as WO2000065062, which  
 CC contains additional sequences as shown in AAB99132-AAB99143,  
 CC AAF75903-AAH75920 and AAG66436.  
 XX  
 SO Sequence 258 AA;  
 Query Match 15.3%; Score 252.5; DB 22; Length 258;  
 Best Local Similarity 29.8%; Pred. No. 4.5e-19;  
 Matches 78; Conservative 57; Mismatches 99; Indels 29; Gaps 10;

OY 55 GETPPIVITSGKGVGKTTTANVGLSLAPYGSVAIDADGLPNDLLGLEKNVY 114  
 DB 4 GLTAATAVIVTGRGAGK-TTTANISTYFAQAGYVLAIGDLPLPGLHPLDN-VKY 61  
 OY 115 TCVEVI-NGCRLDQALVDRKSNFELLCISKPRSKLPMGFGSKALEWLVALKTPPE3 173  
 DB 62 TLHSIVFPMNMPREVAIVYHEQTVY-VMGSPLELDVGSQFLKEVLENLKYK...116  
 OY 174 SPDFIIDCPAGIDAGFTTATTPANENAVLTPPDITLPPADPVYG---LECPGIP---227  
 DB 117 -YPVIFVNSPTGVPPTLPAFESFNYQIVIEIPSPYSFETWENEVLKLYALDOPFY 175  
 OY 228 -DIKMIIVPVPTDMTGYEDMSVLD---VQEMIGSLLVIPEDSEVISTNPGPIV 281  
 DB 176 LDVGVIKVR-----EADVDIKIVETIEEDIGVPGVLPFDVAVESINVGPVL 228  
 OY 282 LNKPTLAGLAFEGQAMPLEQ 303  
 DB 229 VYPPSPALAFYENG-GITER 249

RESULT 13  
 AAY10990  
 ID AAY10990 standard; Protein; 294 AA.  
 XX  
 AC AAY10990;  
 XX  
 DT 08-JUN-1999 (first entry)  
 XX  
 DE H. pylori OFP 06cp11202\_4569693\_c2\_28 cellular protein.  
 XX  
 KW Vaccine; probe; diagnostic; OFP; cell envelope protein;  
 KW secreted protein; cellular protein.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN W09818123-A1.  
 XX  
 PD 07-MAY-1998.  
 XX  
 PF 28-OCT-1997; 97WO-US19575.  
 XX  
 PR 14-JUL-1997; 97US-0891928.

PR 28-OCT-1996; 96US-0739150.  
 PR 06-DEC-1996; 96US-0759739.  
 XX  
 XX (ASTR ) ASTRA AB.  
 PA  
 PI Alm RA, Smith D;  
 XX  
 DR WPI: 1998-271811/24.  
 DR N-PSDB; AAX30457.  
 XX  
 PT Helicobacter pylori nucleic acids and proteins - used to develop  
 PT products for the detection, prevention and treatment of *H. pylori*  
 PT infections  
 XX  
 PS Claims 27, 31, Page 195; 279pp; English.  
 XX  
 CC Recombinant or substantially pure preparations of *H. pylori* polypeptides  
 CC are disclosed, together with the nucleic acids encoding them. In all,  
 CC 73 ORFs are shown. The proteins are variously cell envelope proteins,  
 CC secreted proteins or other cellular proteins. Vaccines containing the  
 CC nucleic acids or proteins are claimed, as are probes containing at least  
 CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful  
 CC for treating or reducing the risk of *H. pylori* infections, and the  
 CC probes can be used diagnostically for detecting the presence of  
 CC Helicobacter in a sample. The products are also of use in screening  
 CC for compounds having the ability to interfere with the *H. pylori* life  
 CC cycle or to inhibit *H. pylori* infection.  
 XX  
 SO Sequence 294 AA;  
 Query Match 13.5%; Score 224; DB 19; Length 294;  
 Best Local Similarity 29.1%; Pred. No. 7.8e-16;  
 Matches 87; Conservative 40; Mismatches 106; Indels 66; Gaps 11;

OY 55 GETPPIVITSGKGVGKTTTANVGLSLAPYGSVAIDADGLPNDLLGLEKNVY 114  
 DB 25 GNT-KFLAITSKGVGKSNISNLAYS-VYGVVGVVFAAG-TLALNLVFTVTHYH 82  
 OY 115 TCVEVI-NGCRLDQALVDRKSNFELLCISKPRSKLPMGFGSKALEWLVALKTPPE3 173  
 DB 93 -LHALGSEALQ-----ETICEIEPGLCLPDSGFEITLYISAEALNPFV 120  
 OY 165 DALKTPREG---SPDFIIDCPAGIDAGFTTATTPANENAVLTPPDITLPPADPVY 221  
 DB 130 D-----EBGVLSLDYVIDTGAGIGATTGAFLNDSQVIVITTPDSALIDA-----Y 178  
 OY 222 ECGGIP-----DIYIVNRY-----PTDMTGYEDMSVLDVGEMLGSLLVYIP 265  
 DB 172 ACIRIKSKYDELEFLIMVAVCPREGATYER-LFYAKYRIAS LELHYLGAIE 231  
 OY 266 EDSEVISTNPGPPLVANKRPPTLAGLAFEGQAMPLEQSMVAVVVEPPKPGFSFF 324  
 DB 232 NSDLKRYVERKILRIARHDLPSQSIDQMSLVSKLETGTL...EIRKGLKSF 286

RESULT 14  
 AAM20738  
 ID AAM20738 standard; protein; 297 AA.  
 XX  
 AC AAM20738,  
 XX  
 DT 16-JUL-1997 (first entry)  
 XX  
 DE H. pylori cytoplasmic protein, 06cp20302orf8.  
 XX  
 KW Cytoplasmic, vaccine, prevention, treatment, infection, identification;  
 KW binding compound; bacterium; life cycle; activator; bactericidal; inhibitor;  
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN W09640893-A1.





Thu Apr 17 10:30:02 2003

us-10-067-989-1.rag

Page 16

Search completed: April 16, 2003, 09:06:09  
Job time : 41 secs

---

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 09:04:32 ; Search time 45 Seconds  
(Without alignments)

Title: US-10-067-989-1  
696,440 Million cell updates/sec

Perfect score: 1654

Sequence: 1 MASRLRSTHNSLLPSSL ..... KATWVEEPKPPSPFSG 326

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	792	47.9	282	T07303	cell division inh
2	657.5	39.8	268	A12237	septum site-determ
3	643	38.9	266	S76970	hypothetical prote
4	572.5	34.7	276	D75478	septum site-determ
5	559.5	33.8	263	B97054	septum site-determ
6	546.5	33.0	266	AB1630	cell division inh
7	545.5	33.0	266	AB1267	cell division inh
8	536.5	32.4	264	C84C38	septum site-determ
9	523	31.6	268	G45239	septum site-determ
10	516	31.2	262	F70375	cell division inh
11	514.5	31.1	269	C82695	septum site-determ
12	506.5	30.6	268	C64561	cell division inh
13	500.5	30.3	268	B71945	cell division inh
14	487	29.4	270	AE0253	septum site-determ
15	483	29.2	270	AH0724	septum site-determ
16	481.5	29.1	271	C98357	septum site-determ
17	481.5	29.1	271	AB2356	cell division inh
18	480.5	29.1	271	F83239	cell division inh
19	479	29.0	270	CCEC1D	cell division inh
20	479	29.0	270	E90837	cell division inh
21	479	29.0	270	F85695	cell division inh
22	475.5	28.7	271	A72300	septum site-determ
23	468	28.3	276	C82136	septum site-determ
24	460	27.8	270	B84968	septum site-determ
25	452	27.3	271	A96C11	probable cell divi
26	449.5	27.2	271	C81230	septum site-determ
27	388	23.5	245	AE3625	cell division inh
28	366	22.1	260	F75175	cell division inh
29	344.5	20.8	259	B69113	cell division inh

30	341.5	20.6	245	2	C71105	probable cell divi
31	335	20.3	263	2	H69336	cell division inh
32	327.5	19.8	245	2	A75056	cell division inh
33	300	18.1	264	1	C64368	cell division inh
34	266	16.1	252	2	F71191	hypothetical prote
35	262.5	15.9	305	2	A84363	cell division inh
36	254	15.4	217	2	H71038	probable cell divi
37	249	15.1	280	2	B83463	flagellar synthe
38	247.5	15.0	251	2	H75150	cell division inh
39	246.5	14.9	288	2	D83954	hypothetical prote
40	243.5	14.7	323	2	B84251	cell division inh
41	242	14.6	263	2	B64321	cell division inh
42	241	14.6	268	2	C81422	probable ATP-bind
43	239.5	14.5	304	2	D71290	probable ATP-bind
44	238.5	14.4	295	2	E70133	ATP-binding protei
45	236	14.3	287	2	D97164	ATPases involved i

#### ALIGNMENTS

##### RESULT 1

T07303

cell division inhibitor - Chlorella vulgaris chloroplast  
C:Species: chloroplast Chlorella vulgaris

C:Date: 14-May-1999 #sequence\_rev: 14-May-1999 #text\_change 21-Jul-2000

C:Accession: T07303

R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakae

Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997

A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlo

A:Reference number: 215985, MID:97030241, PMID:9159184

A:Accession: T07303

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1,282 <MAX>

A:Cross-references: EMBL:AE001684; MID:32224352, FIDR:BA57951.1; FID:32224467

A:Genetics:

A:Gene: mind

A:Superfamily: cell division inhibitor mind

C:Keywords: chloroplast

Query Match 47.9%; Score 792; DB 2; Length 282;

Best Local Similarity 60.4%; Pred. No. 56-56;

Matches 165; Conservative 39; Mismatches 55; Indels 14; Gaps 5;

QY	59	RIVVITSGKGVGKTTTANVGLSLARYGSVAIDADGLRNLDLLLENRVYTCVE	118
DB	17	RIVVITSGKGVGKTTTANVGLSLARYGSVAIDADGLRNLDLLLENRVYTCVE	76
QY	119	VINDGCPILQALVDPDKPNSFELLCTSKPPSKLPMGFGKALEMLVDALKTPPEGSPDI	178
DB	77	IVGGCRIDQALIRDKPKNKLALISKPKQK--YVTRKMCNLTDSVK--ELGFGCV	131
QY	179	IIDPGADAGFITAIPANEAVLVTPDITLADADPYTGLCEGIPDIPMIVNVPYT	238
DB	132	LIDPGADIVGINALASQAEAVITVTPETALRDADRVAGLLENGIIVNKLIVNRVP	191
QY	239	DMIKEDMSVLDVQENLGLSLGVPEDSEVIRSTNRGPIVLNKPPLLAGLAEQAAM	298
DB	192	DMICNDMSVFDVQENLGLSLGVPEDSEVIRSTNRGPIVLNKPPLLAGLAEQAAM	251
QY	299	RIVVITSGKGVGKTTTANVGLSLARYGSVAIDADGLRNLDLLLENRVYTCVE	118
DB	252	PLIKQD---VFIDLTSPQKMPFKLCEFFLG	280

##### RESULT 2

A12237

septum site-determining protein [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_rev: 14-Dec-2001 #text\_change 30-Jun-2002





A:Residues: 1-264 <STO>  
A:Cross-references: GB:AP001517; GB:BA000004; NID:G10175500; PIDN:BA06746.1; GSPDB:GN00  
A:Experimental source: strain C 125  
C:Genetics:  
A:Gene: mind  
C:Superfamily: cell division inhibitor mind

Query Match 32.4%; Score 536.5; DB 2; Length 264;  
Best local similarity 41.1%; Pred. No. 1.4e-35;  
Matches 109; Conservative 66; Mismatches 83; Indels 7; Gaps 3;

OY 61 VITSGKGVGKTTTAVNGSLARVGSVAIDADIGLRLDLLGLERNVYTCVEVI 120  
DB 5 IVITSGKGVGKTTTAVNGSLARVGSVAIDADIGLRLDLLGLERNVYTCVEVI 64  
OY 121 NGDCRLDQALVDRKWSNFELICISPKRSKLPWGFGKALEMLVDAKTRPEGSPDFII 180  
DB 65 EGRCKLQKALIKRKFELCNLLPAQTKD--SAVTEPQMEIVEIKQ---EYDVVL 118  
OY 181 DCPAGIDAGFTTATTPANEAVLVTPPTITLADRDVTGLLECDGIRDIKIVNRVETDM 240  
DB 119 DCPAGIEGFKAAVAGAKAIVVTPETISVYDADRIIGLLEKEVEAPRLVNRIRGHM 178  
OY 241 IKGEMMSVLVDQEMGLSLGVIPEDSEVIRSTNRGFPVLNKPPTLAGLAEQAAMRL 300  
DB 179 MNGEMLVDELVSILALIELGIVDDENVIKFSKGEFIALH-FDSKASVAIRNIARRI 237  
OY 301 VEODSMKAVWEEBPKRGKGFSPFG 325  
DB 238 LGETVPLMSFEQKGVLAIRKISFEG 262

RESULT 9  
cell division inhibitor mind [validated] - Bacillus subtilis  
N:Alternate names: septum placement determinant mind  
C:Species: Bacillus subtilis  
C:Date: 10-Jun-1993 #sequence revision 04-Oct-1996 #text change 19-Jan-2001  
C:Accession: S11205; G45239; F45240; A69658; S27521; S29866  
R:Ref: S. Price, C.W.  
Mol. Microbiol. 7, 601-610, 1993  
A:Title: The mind locus of Bacillus subtilis lacks the mind determinant that provides t  
A:Reference number: S11204; MUID:93213302; PMID:8459776  
A:Accession: S11205  
A:Molecule type: DNA  
A:Residues: 1-268 <LEB>  
A:Cross-references: EMBL:Z15113; NID:G49307; PIDN:CA078818.1; PID:G580893  
R:Levin, P.A.; Marzolis, P.S.; Setlow, P.; Losick, R.; Sun, D.  
J. Bacteriol. 174, 6717-6728, 1992  
A:Title: Identification of Bacillus subtilis genes for septum placement and shape determ  
A:Reference number: A45239; MUID:93015731; PMID:11400224  
A:Accession: G45239  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-268 <LEV>  
A:Cross-references: GB:M06143; NID:G14285; PIDN:AA023401.1; PID:G142859  
R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C:Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capuani, V.; Carter, N.M.; Cho  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Furma, S.; Galluzzi, A.; Gall  
Lich, J.; Harwood, C.P.; Henaux, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koster, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
V.; M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,  
Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadleir, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeder, F.; Scoffone, F.; Sekiguchi, Y.; Skowaska, A.; Serr  
Akenchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Zumbstein, A.  
A:Reference number: A69658; MUID:98044033; PMID:9384377  
A:Accession: A69658  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-268 <KUN>  
A:Cross-references: GB:299218; GB:AL009126; NID:G2C35200; PIDN:CA014753.1; PID:G2C35264  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: mind  
A:Start codon: TTG  
C:Function:  
A:Description: mediates, together with mind, cell division site selection by specifically  
C:Superfamily: cell division inhibitor mind  
C:Keywords: cell division control; nucleotide binding; P-loop  
F10-17/Region: nucleotide-binding motif A (P-loop)  
F16/Binding site: ATP (Lys) #status predicted

Query Match 31.6%; Score 523; DB 1; Length 268;  
Best local similarity 41.1%; Pred. No. 1.7e-34;  
Matches 111; Conservative 68; Mismatches 77; Indels 14; Gaps 7;

OY 61 VITSGKGVGKTTTAVNGSLARVGSVAIDADIGLRLDLLGLERNVYTCVEVI 120  
DB 5 IVITSGKGVGKTTTAVNGSLARVGSVAIDADIGLRLDLLGLERNVYTCVEVI 64  
OY 121 NGDCRLDQALVDRKWSNFELICISPKRSKLPWGFGKALEMLVDAKTRPEGSPDFII 180  
DB 65 EGRCKMHALVKDRFPDL-LYLMPPAQSDKTAVAPEQIKNVQGLKQ---EYDVVL 119  
OY 181 DCPAGIDAGFTTATTPANEAVLVTPPTITLADRDVTGLLECDGIRDIKIVNRVETDM 240  
DB 120 DCPAGIEGKYNAVAGADKAIVVTPETISVYDADRIIGLLEKEVEAPRLVNRIRGHM 179  
OY 240 MIKEDMMMSVLVDQEMGLSLGVIPEDSEVIRSTNRGFPVLNKPPTLAGLAEQAAMRL 299  
DB 180 LKMGDMDIDELVCHSLIDLLGIVADDEVIRKASNGERIAMD-PKNASTAYRNINAP 238  
OY 300 LVEDSMKAVWEEBPKRGKGF-----SFG 325  
DB 239 IL-GEVPLQVLEQNK--GMKAKISFEG 265

RESULT 10  
cell division inhibitor (septum placement) mind2 - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence revision 08-May-1998 #text change 26-Aug-1999  
A:Accession: F70375  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lennox, A.L.; Graham, D.E.; Ove  
V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: F70375  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-262 <ADP>  
A:Cross-references: GB:AE0090712; NID:G2983411; PIDN:AA006906.1; PID:G2983411; GP:AF00065;  
A:Experimental source: strain VFS  
C:Genetics:  
A:Gene: mind2  
C:Superfamily: cell division inhibitor mind

Query Match 31.2%; Score 516; DB 2; Length 262;  
Best local similarity 41.4%; Pred. No. 6e-34;

Matches 108; Conservative 65; Mismatches 70; Indels 18; Gaps 3;

QY 60 IVITSGKGVGKTTTANVGLSLARYGFSVAIDADLGRNLDLLGLENNVNTCEV 119  
 Db 4 VVITSGKGVGKTTTANVGLSLARYGFSVAIDADLGRNLDLLGLENNVNTCEV 63

QY 120 INGDCRLDQALVPRDPMWGNFELLCTSKPRSKLPMFGFGKALEMVLDAKTRPESDPFI 179  
 Db 64 LEGRPYREKALVKDRKGLSLMLLPANQRANKVDIDIE---KMKRVEELKNSGNDYIL 119

QY 180 IDCPAGIDAGITATTPANEAULTPTDITATLRDADRVYGLLECDGIPDIKMIYNEVPTD 239  
 Db 120 VDSPIAGEKGFQIAVSPADKALIVNEPEVSSIPDADPVYIGLESEMERNKXVITNPKME 179

QY 240 MIKGDPMKSVLDVQEMGLSLGVIPDESEVIRSTNGPFLVNLK--PPLTAGLAEQAA 297  
 Db 180 MYKRGAMLSVEIDVILAEFIIIGIIPERPKLVDFETNGEPYVLEPPASCAIT.... 233

QY 298 WRLVEODSKAMVVEEPKRR 318  
 Db 234 -----DTRAPRLMGESTLPR 248

RESULT 11  
 C82695  
 septum site-determining protein Xfl321 [imported] - Xylella fastidiosa (strain 985C)  
 C/Species: Xylella fastidiosa  
 C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
 C/Accession: C82695  
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
 Nature 406, 151-157, 2000  
 A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A/Reference number: A82515; PMID:20365717, PMID:10910347  
 A/Note: for a complete list of authors see reference number A59328 below  
 A/Accession: C82695  
 A/Status: preliminary  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-269 <SIM>  
 A/Cross-references: GB:AE003965; GB:AE003849; NID:G9106313; PIND:AAF8130.1; GSPDB:GN001  
 A/Experimental source: strain 985C  
 R:Stimpson, A.J.G.; Peimach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
 Brites, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H  
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Frana, S.C.; Franco, M.C.; Frohm  
 J.D.; Junqueira, M.L.; Kemper, F.L.; Kitejima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.P.; Martins, C.L.; Marques, M.V.; Martins, E  
 A:Authors: Martins, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Miyaki, C.Y.;  
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
 Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
 A:Authors: da Silva, A.C.P.; da Silva, F.P.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L...  
 A/Reference number: A59328  
 A/Contents: annotation  
 C/Genetics:  
 A:Gene: Xfl321  
 C/Superfamily: cell division inhibitor mind

Query Match 31.1%; Score 514.5; DB 2; Length 269;  
 Best Local Similarity 43.5%; Pred. No. 8; Ce-34;  
 Matches 120; Conservative 58; Mismatches 79; Indels 19; Gaps 8;

QY 60 IVITSGKGVGKTTTANVGLSLARYGFSVAIDADLGRNLDLLGLENNVNTCEV 119  
 Db 4 VVITSGKGVGKTTTANVGLSLARYGFSVAIDADLGRNLDLLGLENNVNTCEV 63

QY 120 INGDCRLDQALVPRDPMWGNFELLCTSKPRSKLPMFGFGKALEMVLDAKTRPESDPFI 179  
 Db 64 IDGEATLQALIKDPDNILAAQTPDKDAITREGE--VEVNLNELQ--EGF-DYIC 118

QY 180 IDCPAGIDAGITATTPANEAULTPTDITATLRDADRVYGLLECDGIPDIKMIYNEVPTD 239  
 Db 119 CDSPIAGEKGFQIAVSPADKALIVNEPEVSSIPDADPVYIGLESEMERNKXVITNPKME 178

QY 233 VNRVDTMIKGDPMKSVLDVQEMGLSLGVIPDESEVIRSTNGPFLVNLKPTIAGLA 292  
 Db 179 LTRSPARVESGEMLSIADVEEVLGKALGVIPESGCVLNASKKEPVLDN-NSLAGLA 237

QY 293 FEQAMRLVEOD-SMKAMVEEPPKRRGFS-FFGG 326  
 Db 238 YEDAVGPIGCEHMPPTVE---YKGFESKLCFG 269

RESULT 12  
 C64561  
 cell division inhibitor mind Helicobacter pylori (strain 26695)  
 C/Species: Helicobacter pylori  
 C/Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 02-Feb-2001  
 C/Accession: C64561  
 R:Toth, J.F.; White, O.; Kariya, A.F.; Clayton, R.A.; Sullivan, G.J.; Fleischmann, R.D.;  
 Peterson, S.; Loftus, B.; Richardson, P.; Dodson, P.; Khajak, H.G.; Giron, A.; McInney  
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weisman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.  
 Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karik, P.D.; Smith, H.O.; Fraser, C.J  
 Nature 388, 539-547, 1997  
 A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A/Reference number: A64520; PMID:97394467; PMID:9252185  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-268 <DOM>  
 A/Cross-references: GB:AE000551; GB:AF130511; NID:G331433; PIND:AAU07400.1; FID:G2313432  
 C/Genetics:  
 A:Gene: mind  
 C/Function:  
 A/Description: an ATPase required for correct placement of cell division site  
 C/Superfamily: cell division inhibitor mind  
 C/Keywords: ATP; cell division control; nucleotide binding; P-loop  
 F10-17/Region: nucleotide-binding motif A (P-loop)  
 F16/Binding site: ATP (Lys) #status predicted

Query Match 30.6%; Score 506.5; DB 2; Length 268;  
 Best Local Similarity 40.1%; Pred. No. 3; Ce-33;  
 Matches 110; Conservative 65; Mismatches 80; Indels 19; Gaps 5;

QY 60 IVITSGKGVGKTTTANVGLSLARYGFSVAIDADLGRNLDLLGLENNVNTCEV 119  
 Db 4 VVITSGKGVGKTTTANVGLSLARYGFSVAIDADLGRNLDLLGLENNVNTCEV 63

QY 120 INSDCPDQALVPRDPMWGNFELLCTSKPRSKLPMFGFGKALEMVLDAKTRPESDPFI 179  
 Db 64 MEKCNLSQALITDKTKNLSPLASQSKNIDREKVAI--LINALR---ADFDYIL 117

QY 180 IDCPAGIDAGITATTPANEAULTPTDITATLRDADRVYGLLECDGIPDIK-----MI 232  
 Db 118 IDSPAGIESGEHALIADALVALVTPPEVSSLPDSPPVVGIIIDAKSNPAKMEVHMLI 177

QY 233 VNRVDTMIKGDPMKSVLDVQEMGLSLGVIPDESEVIRSTNGPFLVNLKPTIAGLA 292  
 Db 178 INRLKPELVANGEMISIEEVLKILCLPIIGIIPEDHIIISAINKEPVI--RTDESAKA 235

QY 293 FEQAMRLVEODSKAMVVEEPPKRRGFSFGG 326  
 Db 236 YQRTPRILGEE---VEYEFKPKGFPSALKG 265

RESULT 13  
 B71945  
 cell division inhibitor - Helicobacter pylori (strain 199)  
 C/Species: Helicobacter pylori  
 A/Variety: strain 199  
 C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 20-Aug-1999  
 C/Accession: B71945  
 R:Alm, P.A.; Ling, L.S.L.; Moir, D.T.; King, P.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
 Jves, C.; Gibson, R.; Weberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; J  
 Nature 397, 176-180, 1999  
 A/Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71900; MUID:99129547; PMID:9933682  
 A:Accession: B71945  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-268 <PAR>  
 A:Cross references: GB:AF001468; GB:AF001439; NID:G4154818; PICH:AA09995.1; PID:G415485  
 A:Experimental source: strain 109  
 C:Genetics:  
 A:Gene: mind  
 C:Superfamily: cell division inhibitor mind

Query Match 30.3%; Score 500.5; DB 2; Length 268;  
 Best Local Similarity 39.8%; Pred. No. 1,1e-32;  
 Matches 109; Conservative 65; Mismatches 81; Indels 19; Gaps 5;

DB 60 IVVTSGGKGVGKTTTANVGLSLAFYGSVAIDADGLRNDLLGLENNVNTCWE 119  
 4 VVTISGKGVGKSTTTANLAIGLAEKGVAVDPDGLRNDLLGLENNVNTCWE 63  
 DB 120 TNGDRLDQALVPRKMSNFELLCISKRSKLPWGFGKALEMVDALKTRPGSPDFI 179  
 64 MERNLNLSCALITPKTKNLSPLASGSKNDLDEKVAL-LINLR ADEPDYIL 117  
 DB 180 IDCPAGIDAGFTTATTPANEAVLVTPDITARDADRVTCGLECGIR.... DIKMI 232  
 118 IDSPAGISGEFHAILHADMALVVTPEVSSLRSDRVIGIIDAKSRASGEVHKHLI 177  
 DB 233 VNPVTNMIKEDMMSVLDVQEMGLSLGVIPEDSEVIRSTNKGFLVLPPTIAGLA 292  
 DB 178 INFRPELVANGEMISIEVVKILCLPLIGIPEDHIIISATNKGPRVI..PIDCESAFA 235  
 DB 293 FEQAMRLVEQDSMKAVWVEEPPKRGFFSFGG 326  
 DB 236 VQPIPTPLUGEE---VFVVEFKVKGFFSALKG 265

# RESULT 14

AE0253  
 septum site-determining protein [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis  
 C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001  
 C:Accession: AE0253  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.P.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, P.M.; Dougan, G.; H., M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett, N.; Nature 413, 523-527, 2001  
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A:Reference number: AE0001; MUID:21470413; PMID:11586360  
 A:Accession: AE0253  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-270 <KIP>  
 A:Cross references: GB:AL533862; PICH:CA05498.1; PID:G1533862; GSPDR:GN00175  
 C:Genetics:  
 A:Gene: mind  
 C:Superfamily: cell division inhibitor mind

Query Match 29.4%; Score 487; DB 2; Length 270;  
 Best Local Similarity 41.2%; Pred. No. 1,3e-31;  
 Matches 114; Conservative 55; Mismatches 90; Indels 18; Gaps 7;

DB 59 RIVVTSGKGVGKTTTANVGLSLAFYGSVAIDADGLRNDLLGLENNVNTCWE 118  
 3 RIVVTSGKGVGKTTTANVGLSLAFYGSVAIDADGLRNDLLGLENNVNTCWE 62  
 DB 119 VINGDRLDQALVPRKMSNFELLCISKRSKLPWGFGKALEMVDALKTRPGSPDFI 178  
 63 VINGDRLDQALVPRKMSNFELLCISKRSKLPWGFGKALEMVDALKTRPGSPDFI 117  
 DB 179 IIDCPAGIDAGFTTATTPANEAVLVTPDITARDADRVTCGLECGIR.... DIK-- 230  
 118 IIDCPAGIDAGFTTATTPANEAVLVTPDITARDADRVTCGLECGIR.... DIK-- 177

DB 231 MIVNRVTMIKEDMMSVLDVQEMGLSLGVIPEDSEVIRSTNKGFLVLPPTIAG 290  
 DB 178 LLTRVYNGRVNKGDMUSMELVDILPLVAVIPEDQSVLPASNGEVLIDKESD-AG 236  
 DB 291 LAFCQAMRLVEQDSMKAVWVEEPPKRGFFS-FFGG 326  
 DB 237 KAYADVPLUGEE-PPPFVEE--KGFPLPLFQG 270

# RESULT 15

AH0724  
 septum site-determining protein [imported] - Salmonella enterica subsp. enterica serovar

C:Species: Salmonella enterica subsp. enterica serovar Typhi  
 A:Note: this species has also been called Salmonella typhi  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AH0724  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.P.; Pickard, D.; Main, J.; Churcher, C.; Connor, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Mout, S.; O'Garra, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Park, C.; Cronin, M.; Fetherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar  
 A:Reference number: AB0502; PMID:11677608  
 A:Accession: AH0724  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-270 <PAR>  
 A:Cross references: GB:AL533862; PICH:CA05498.1; PID:G1533862; GSPDR:GN00176  
 C:Genetics:  
 A:Gene: sty1945  
 C:Superfamily: cell division inhibitor mind

Query Match 29.3%; Score 483; DB 2; Length 270;  
 Best Local Similarity 41.2%; Pred. No. 2,8e-31;  
 Matches 114; Conservative 57; Mismatches 88; Indels 18; Gaps 8;

DB 59 RIVVTSGKGVGKTTTANVGLSLAFYGSVAIDADGLRNDLLGLENNVNTCWE 118  
 3 RIVVTSGKGVGKTTTANVGLSLAFYGSVAIDADGLRNDLLGLENNVNTCWE 62  
 DB 119 VINGDRLDQALVPRKMSNFELLCISKRSKLPWGFGKALEMVDALKTRPGSPDFI 178  
 63 VINGDRLDQALVPRKMSNFELLCISKRSKLPWGFGKALEMVDALKTRPGSPDFI 117  
 DB 179 IIDCPAGIDAGFTTATTPANEAVLVTPDITARDADRVTCGLECGIR.... DIK-- 230  
 DB 118 VCDSPAGIDAGFTTATTPANEAVLVTPDITARDADRVTCGLECGIR.... DIK-- 177  
 DB 231 MIVNRVTMIKEDMMSVLDVQEMGLSLGVIPEDSEVIRSTNKGFLVLPPTIAG 290  
 DB 178 LLTRVYNGRVNKGDMUSMELVDILPLVAVIPEDQSVLPASNGEVLIDKESD-AG 236  
 DB 291 LAFCQAMRLVEQDSMKAVWVEEPPKRGFFS-FFGG 326  
 DB 237 KAYADVPLUGEE-PPPFVEE--KGFPLPLFQG 270

Search completed: April 16, 2003, 09:08:36  
 Job time : 46 secs



GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd

OM protein - protein search, using sw model

Run on: April 16, 2003, 09:03:53 ; Search time 88 Seconds  
(without alignments)  
763.311 Million cell updates/sec

Title: US-10-067-989-1  
Perfect score: 1654  
Sequence: 1 MASLRJFSTHQSLPLPSL KAWWEPRKPKRPFPPFGS 326

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1654	100.0	326	10 Q9MBA2	Q9MBA2 arabidopsis
2	1175.5	71.1	295	10 Q9PV40	Q9PV40 tagetes ere
3	1067.5	64.5	306	10 Q9LMY6	Q9LMY6 oryza sativ
4	786	45.7	274	8 Q9T3P6	Q9T3P6 nephrolepis
5	752	45.5	359	8 Q9TUB6	Q9TUB6 prototheca
6	657.5	39.8	268	16 Q8YRJ0	Q8YRJ0 anabaena sp
7	606.5	36.7	264	16 Q8RQV1	Q8RQV1 fuscharteri
8	574.5	34.7	264	16 Q8RBB9	Q8RBB9 thermomater
9	573.5	34.7	276	16 Q8RMB7	Q8RMB7 thermomater
10	563	34.0	265	16 Q8X1I1	Q8X1I1 clostridium
11	559.5	33.8	263	16 Q970M4	Q970M4 clostridium
12	546.5	33.0	266	16 Q92BH0	Q92BH0 listeria in
13	545.5	32.4	266	16 Q8Y6Y7	Q8Y6Y7 listeria mo
14	536.5	32.4	264	16 Q9K8H9	Q9K8H9 bacillus ha
15	516	31.2	262	16 Q670J3	Q670J3 aquifex ae
16	514.5	31.1	269	16 Q9PD08	Q9PD08 xylella fas

17	487	29.4	270	16 Q8ZSE6	Q8ZSE6 yersinia pe
18	483	29.2	270	16 Q8ZPI0	Q8ZPI0 salmonella
19	483	29.1	270	16 Q8Z680	Q8Z680 salmonella
20	481.5	29.1	271	16 Q8UAW9	Q8UAW9 agrobacteri
21	480.5	29.1	271	16 Q8YXZ6	Q8YXZ6 pseudomonas
22	479	29.0	179	10 Q9SPD0	Q9SPD0 oryza sativ
23	468	28.3	276	16 Q8KON8	Q8KON8 vibrio chol
24	461.5	27.9	271	16 Q8U0G9	Q8U0G9 rhizobium i
25	452	27.3	271	16 Q9ZT20	Q9ZT20 rhizobium m
26	449.5	27.2	271	2 Q9AG19	Q9AG19 neisseria g
27	449.5	27.2	271	16 Q9U0Y6	Q9U0Y6 neisseria m
28	440	26.6	271	16 Q8XU29	Q8XU29 ralscrista s
29	388	23.5	245	16 Q8YBH4	Q8YBH4 bruceella me
30	380	23.0	264	17 Q8U0H8	Q8U0H8 pyrococcus
31	366	22.1	260	17 Q9V165	Q9V165 pyrococcus
32	344.5	20.8	259	17 Q27868	Q27868 methanobact
33	341.5	20.6	245	17 Q58346	Q58346 pyrococcus
34	335	20.3	263	17 Q29562	Q29562 archaeoglob
35	327.5	19.8	245	17 Q8UY80	Q8UY80 pyrococcus
36	314.5	19.0	245	17 Q8U311	Q8U311 pyrococcus
37	307.5	18.6	269	17 Q8TVS4	Q8TVS4 methanopyru
38	266	16.1	252	17 Q59470	Q59470 pyrococcus
39	262.5	15.9	305	17 Q9HNF1	Q9HNF1 halobacteri
40	254	15.4	217	17 Q59249	Q59249 pyrococcus
41	250	15.1	319	2 Q9L775	Q9L775 xanthomonas
42	249	15.1	280	16 Q9KRC0	Q9KRC0 pseudomonas
43	248.5	15.0	253	17 Q8TXK6	Q8TXK6 pyrococcus
44	247.5	15.0	251	17 Q9V1R2	Q9V1R2 pyrococcus
45	246.5	14.9	288	16 Q9KAS4	Q9KAS4 bacillus ha

## ALIGNMENTS

RESULT 1

Q9MBA2 PRETMINARY, FFT, 326 AA.

ID Q9MBA2

AC Q9MBA2: 01-OCT-2000 (TREMBLrel, 15, Created)

DT 01-OCT-2000 (TREMBLrel, 15, Last sequence update)

DT 01-MAR-2002 (TREMBLrel, 20, Last annotation update)

DE Mind (Septum site-determining Mind).

GN MIND.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;

OC Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, Rosidae;

OC euroids II, Brassicales, Brassicaceae, Arabidopsis.

OX NCBI\_TaxID=3702;

OX [1]

RP SEQUENCE FROM N.A.

RA Kim M., Fujiwara M., Kanamaru K., Tanaka K., Takahashi H.,

RT "Arabidopsis thaliana mind homolog involved in plastid division";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=COLUMBIA;

RX MEDLINE=98290546; PubMed=9628582;

RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,

RA Tabata S.,

RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.

RT Sequence features of the regions of 1,456,315 bp covered by nineteen

RT physically assigned pl and TAC clones."

RL DNA Res. 5, 41-54(1998).

RL EMBL, AB030278, BAA92263.1, -

DR EMBL, AB009056, BAB08725.1, -

DR InterPro, IPR007077, ATPase\_Para.

DR Pfam, PF00991, Para, 1.

SQ SEQUENCE 326 AA; 35690 MW; BE7AEPA307167825 CRC64;

Query Match 100.0%; Score 1654; DB 10; Length 326;  
Best Local Similarity 100.0%; Pred. No. 6,4e-128;  
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MASLRFSTNHOSLLPSSLSOKTILSSPRVNNPSRSPIRSVLQNRKRELAGETPRT 60
DB 1 MASLRFSTNHOSLLPSSLSOKTILSSPRVNNPSRSPIRSVLQNRKRELAGETPRT 60
QY 61 VVITSGKGVGKTTTAAVNGSLARVGSVAIDADLGLPDLGLLENRVNTTCEVI 120
DB 61 VVITSGKGVGKTTTAAVNGSLARVGSVAIDADLGLPDLGLLENRVNTTCEVI 120
QY 121 NGDCGLDQALVDPKWSNFEELCISKPSKLPNGHGALEMLVDALKTREPSGPDFI 180
DB 121 NGDCGLDQALVDPKWSNFEELCISKPSKLPNGHGALEMLVDALKTREPSGPDFI 180
QY 181 DCPAGIDAGFITAITPANAVLVTPTDITLAPNDPVYGLGCGIIRIDIMVYVVTDM 240
DB 181 DCPAGIDAGFITAITPANAVLVTPTDITLAPNDPVYGLGCGIIRIDIMVYVVTDM 240
QY 241 IKGEDMSVLDVQEMLCGLLGVIPEDSEVIRSTNRGPFVLNKPPTLACIAGFAQAMRL 300
DB 241 IKGEDMSVLDVQEMLCGLLGVIPEDSEVIRSTNRGPFVLNKPPTLACIAGFAQAMRL 300
QY 301 VEODSMKAVMVEEPPKKGFFSFFGG 326
DB 301 VEODSMKAVMVEEPPKKGFFSFFGG 326

```

## RESULT 2

```

QY 09FV40 PRELIMINARY; PRT; 295 AA.
AC 09FV40;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE MIND.
OS Tagetes erecta (African marigold).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC Heliantheae; Tagetes.
OX NCBI_TaxID=13708;
RN [1]
RP SEQUENCE FROM N A.
RA Moens C.P., Tian L., Dellapenna D.;
RT "Analysis of carotenoid biosynthetic gene expression during marigold
RT petal development."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF25103.9; AAC10431.1;
DR InterPro: IPR000707; Arpase_Para.
DR Pfam: PF00991; Para; 1.
SQ SEQUENCE 295 AA; 32756 MW; 212BD476BF8747F CRC64;

```

Query Match 71.1%; Score 1175.5; DB 10; Length 295;  
 Best Local Similarity 72.0%; Pred. No. 1.2e-88;  
 Matches 247; Conservative 15; Mismatches 16; Indels 65; Gaps 7;

```

QY 1 MASLRFSTNHOSLLPSSLSOKTILSSPRVNNPSRSPIRSVLQNRKRELAGETPRT 60
DB 1 MASLRFSTNHOSLLPSSLSOKTILSSPRVNNPSRSPIRSVLQNRKRELAGETPRT 60
QY 47 FRRKPELAGETPRIIVITSGKGVGKTTTAAVNGSLARVGSVAIDADLGLPDLGL 106
DB 47 FRRKPELAGETPRIIVITSGKGVGKTTTAAVNGSLARVGSVAIDADLGLPDLGL 106
QY 107 GLENRVNTTCEVINGDCRLDQALVDPKWSNFEELCISKPSKLPNGHGALEMLVDA 166
DB 107 GLENRVNTTCEVINGDCRLDQALVDPKWSNFEELCISKPSKLPNGHGALEMLVDA 166
QY 167 LKTRREGSDPFIIDCPAGIDAGFITAITPANAVLVTPTDITLAPNDPVYGLGCGI 226
DB 167 LKTRREGSDPFIIDCPAGIDAGFITAITPANAVLVTPTDITLAPNDPVYGLGCGI 226
QY 227 RDIKMIIVNRVPTDMLKGEDMSVLDVQEMLCGLLGVIPEDSEVIRSTNRGPFVLN 283
DB 227 RDIKMIIVNRVPTDMLKGEDMSVLDVQEMLCGLLGVIPEDSEVIRSTNRGPFVLN 283

```

```

QY 284 KPPTLAGIAFEQAAAMRLVEDDSMKAVMVEEPPKKGFFSFFGG 326
DB 284 KPPTLAGIAFEQAAAMRLVEDDSMKAVMVEEPPKKGFFSFFGG 326

```

## RESULT 3

```

QY 09LWY6 PRELIMINARY; PRT; 306 AA.
AC 09LWY6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Similar to Chlorella vulgaris C-27 chloroplast DNA.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N A.
RA STRAIN=CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone: P0644B06."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001129; BAA90628.1;
DR InterPro: IPR000707; Arpase_Para.
DR Pfam: PF00991; Para; 1.
SQ SEQUENCE 306 AA; 32341 MW; CECAB38219512690 CRC64;

```

Query Match 64.5%; Score 1067.5; DB 10; Length 306;  
 Best Local Similarity 68.4%; Pred. No. 9.7e-80;  
 Matches 216; Conservative 31; Mismatches 50; Indels 19; Gaps 5;

```

QY 14 LILPSSLSOKTILSSPRVNNPSRSPIRSVLQNRKRELAGETPRTIVITSGKGVGKT 73
DB 14 LILPSSLSOKTILSSPRVNNPSRSPIRSVLQNRKRELAGETPRTIVITSGKGVGKT 73
QY 74 TTTANVGLSLARVGSVAIDADLGLPDLGLLENRVNTTCEVINGDCRLDQALVDP 133
DB 74 TTTANVGLSLARVGSVAIDADLGLPDLGLLENRVNTTCEVINGDCRLDQALVDP 133
QY 134 KRWNSFEELCISKPSKLPNGHGALEMLVDALKTREPSGPDFIIDCPAGIDAGFIT 193
DB 134 KRWNSFEELCISKPSKLPNGHGALEMLVDALKTREPSGPDFIIDCPAGIDAGFIT 193
QY 194 ITPANAVLVTPTDITLAPNDPVYGLGCGIIRIDIMVYVVTDMIKGEDMSVLDVQ 253
DB 194 ITPANAVLVTPTDITLAPNDPVYGLGCGIIRIDIMVYVVTDMIKGEDMSVLDVQ 253
QY 254 EMLGSLGVIPEDSEVIRSTNRGPFVLNKPPTLACIAGFAQAMRLVEDDSMKAVMVEE 313
DB 254 EMLGSLGVIPEDSEVIRSTNRGPFVLNKPPTLACIAGFAQAMRLVEDDSMKAVMVEE 313
QY 314 E-PKRR-GFFSFFGG 326
DB 314 E-PKRR-GFFSFFGG 326

```

## RESULT 4

```

QY 09T3P6 PRELIMINARY; PRT; 274 AA.
AC 09T3P6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Septum-site determining protein.
OS MIND.
OS Nephroselmis olivacea.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;
OC Chlorodendraceae; Chlorodendraceae; Nephroselmis.

```



QY 296 AAMRLVEODSMK 307  
DB 234 IARRL-EGDIVE 244

RESULT 7  
Q8RGV1 PPELIMINAPY; PRT; 264 AA.  
AC Q8RGV1; 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DE Cell division inhibitor MIND.  
GN FN0176  
OS Fusobacterium nucleatum (subsp. nucleatum).  
OC Bacteria; Fusobacteria; Fusobacterium.  
OX NCBI\_TaxID=76856;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 25586;  
PX MEDLINE=21986394; PubMed=1189109;  
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,  
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,  
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,  
RA Larsen N., D'Souza M., Malinas T., Pusch G., Haselkorn R.,  
RA Forstern M., Kyprides N., Overbeek R.,  
RT "Genome sequence and analysis of the oral bacterium Fusobacterium  
nucleatum strain ATCC 25586."  
PL J. Bacteriol. 184:2005-2018 (2002).  
DR EMBL, AE010531; AAL94382.1; -  
KW Complete proteome.  
SQ SEQUENCE 264 AA; 28675 MM; DA74D05DF8ED1F68 CRC64;

Query Match 36.7%; Score 606.5; DB 16; Length 264;  
Best Local Similarity 50.0%; Pred. No. 6,38-42;  
Matches 121; Conservative 55; Mismatches 59; Indels 7; Gaps 3;

QY 59 RIVVITSGKGVGKTTTANVGLSLARYGFSVAIDALGRNLDLLGLENNYTCVE 118  
DB 4 RIVVITSGKGVGKTTTANVGLSLARYGFSVAIDALGRNLDLLGLENNYTCVE 63  
QY 119 VINGCPRLQALVPRKPSNFELLCTSPKPSYPMGFGKALEMLVDALKTRECSDFI 178  
DB 64 VIEEPRLISQAFIRKPKPCNVLVLPAAQIPKNDV--TPGKMSLIDSLK---ASFDYI 117  
QY 179 IIDCPAGIDAGRTATTPANEAVLTPPTDIALRDADRVTCLECGDIRDKIMVNRVT 238  
DB 118 LVDCPAGIEQGFKAIVADEAVVVTTEVSATRDADRIIGLEASGKEPRIVINRLKI 177  
QY 239 PMIKGDMMSVLDOEMGLSLGVIPEDSEVIRSTNGFPPLVINKPPTLAGIAFEQAAW 298  
DB 178 DMVKKKMLSVEDIIIDIGIKLGLVPPDEVISTNGEPLV-KGDSLAAKAFKNIAN 236  
QY 299 RL 300  
DB 237 RI 238

RESULT 8  
Q8RBB9 PPELIMINAPY; PRT; 264 AA.  
AC Q8RBB9; 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DE Septum formation inhibitor-activating ATPase.  
GN MIND OR ITB0903.  
OS Thermomicrobacter tengcongensis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
OC Thermomicrobacteriales; Thermomicrobacteriaceae; Thermomicrobacter.  
X NCBI\_TaxID=119072;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=MB4T / JCM11007;  
RX MEDLINE=21992816; PubMed=11997336;  
RA Bao Q., Tian Y., Li W., Xu Z., Hu S., Dong W., Yang J.,  
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
RT "A complete sequence of *T. tengcongensis* genome."  
PL Genome Res. 12:689-700 (2002).  
DR EMBL, AE013057; AAM24159.1; -  
KW Complete proteome.  
SQ SEQUENCE 264 AA; 29357 MM; 3F85B51D77C6859C CRC64;

Query Match 34.7%; Score 574.5; DB 16; Length 264;  
Best Local Similarity 46.3%; Pred. No. 2,78-39;  
Matches 113; Conservative 56; Mismatches 68; Indels 7; Gaps 3;

QY 61 VVITSGKGVGKTTTANVGLSLARYGFSVAIDALGRNLDLLGLENNYTCVEI 120  
DB 5 VVITSGKGVGKTTTANVGLSLARYGFSVAIDALGRNLDLLGLENNYTCVEI 64  
QY 121 NGDCRLDQALVPRKPSNFELLCTSPKPSYPMGFGKALEMLVDALKTRECSDFI 180  
DB 65 EGQCPKQALIPDKPFTVLVLPAAQTRKTA--TPGKMPKLIQTLK---FNVYIIV 119  
QY 181 DCPAGIDAGRTATTPANEAVLTPPTDIALRDADRVTCLECGDIRDKIMVNRVT 240  
DB 119 DCPAGIEGFFNAIAGADFAIVTTPVSATRDADRIIGLEAELNPLVIRPKKM 178  
QY 241 IKGEDMMSVLDOEMGLSLGVIPEDSEVIRSTNGFPPLVINKPPTLAGIAFEQAAW 300  
DB 119 VKRGMDIEDIIDLIALDLGLVPPDEVISTNGEPLV-KSLASQAVNIVEF 237  
QY 301 VEQD 304  
DB 238 LDRN 241

RESULT 9  
Q8RMB7 PPELIMINAPY; PRT; 276 AA.  
AC Q8RMB7; 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE Septum site-determining protein.  
GN DR0752.  
OS Deinococcus radiodurans.  
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
OC Deinococcaceae; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RI;  
RX MEDLINE=20036896; PubMed=10567266;  
RA White O., Eissen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson P.L.,  
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Salewski C.,  
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann P.D.,  
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RA Fraser C.M.;  
RT "Genome sequence of the radioreistant bacterium Deinococcus  
radiodurans RI."  
PL Science 286:1571-1577 (1999).  
DR EMBL, AE001931; AAF10311.1; -  
DR TIGR, DR0752; -  
DR InterPro: IPR000707; ATPase\_Para.  
DR InterPro: IPR000392; NitrogenaseII.  
DR Pfam: PF00142; fer4\_NiFe; 1.  
DR Pfam: PF00991; Para\_1.  
KW Complete proteome.  
SQ SEQUENCE 276 AA; 29420 MM; AD74FDCFA4582D0C CRC64;

Query Match 34.7%; Score 573 5; DB 16; Length 276;  
 Best Local Similarity 44.3%; Pred. No. 3.5e-39;  
 Matches 125; Conservative 52; Mismatches 72; Indels 33; Gaps 5;

59 RIVVITSGKGVGKTTTANVGLSLARYGSVAIDADLGLRNLDLLGLERNVYTCVE 118  
 12 KVIIVTSGKGVGKTTTANVGLSLARYGSVAIDADLGLRNLDLLGLERNVYTCVE 71  
 119 VINGCPRIDQALVPRKPSNFELLCISPRSKLPMGFGKALEMLVDALKTR----- 170  
 72 VLEGCRKMQALIRKRYENHLRLPASQTRDK-----DALPEVEKEVVGK 117  
 171 --PESSPDEIITDCPAGIDAGFITAITPANEAVALTTPDITALRDADRVGLGECGIDP 228  
 118 LLEEEGFDVILDSFAGIESGRTAAAPAGALVNVNPESSVRDADRIIGLEAQOITE 177  
 229 IPMIVNPTVMKCEGEMSVLDVQEMLSLGVIPEDSEVIRSTRGPFVLYNKPPTL 288  
 178 IRLVVRRLRPKVVASGNMILSIDMDVILGVKPIGIVPEDEGIYSTNVGEPAVLGK--TK 235

289 AGLAPQAMRLVEDQSKAMVVEREPKRGFS----FRG 326  
 236 AGLAPQAMRLVEDQSKAMVVEREPKRGFS----FRG 326

## RESULT 10

08X111 PRELIMINARY; PRT; 265 AA.  
 08X111  
 DT 01-MAR-2002 (Tremblrel 20, Created)  
 DT 01-MAR-2002 (Tremblrel 20, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel 21, Last annotation update)  
 DE Septum site-determining protein.  
 GN MIND OR CPE2138.  
 OS Clostridium perfringens.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Clostridiales; Clostridiaceae; Clostridium.  
 OX NCBI\_TaxID=1502;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=13 / TYPE A;  
 RX PubMed=11792842;  
 RA Shimizu T., Ohtani K., Hisekawa H., Ohshima K., Yamashita A.,  
 RA Shiba T., Ogasawara N., Hattori M., Kunara S., Hayashi H.;  
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic  
 RT flesh-eater".  
 RI Proc Natl Acad Sci U S A. 99:996-1001(2002)  
 DR EMBL, AF003193; BAB1844.1;  
 DR InterPro: IPR000707; ATPase\_Para.  
 DR Pfam: PF00991; Para; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 265 AA; 29086 MW; 23096B312D068DDE CRC64;

Query Match 34.0%; Score 563; DB 16; Length 265;  
 Best Local Similarity 45.0%; Pred. No. 2.4e-38;  
 Matches 118; Conservative 55; Mismatches 79; Indels 10; Gaps 4;

61 VITSGKGVGKTTTANVGLSLARYGSVAIDADLGLRNLDLLGLERNVYTCVE 120  
 5 VITSGKGVGKTTTANVGLSLARYGSVAIDADLGLRNLDLLGLERNVYTCVE 64  
 121 NGDCRIDQALVPRKPSNFELLCISPRSKLPMGFGKALEMLVDALKTRPESSPDEI 180  
 65 ENRCRTKQALIRKRYENHLRLPASQTRDK-----DALPEVEKEVVGK 117  
 181 DCPAGIDAGFITAITPANEAVALTTPDITALRDADRVGLGECGIDRIKMIIVNPTVM 240  
 119 DCPAGIEGFEFNAIVGADRAIVNPEVTSVRDADRVIGKIDAKGIEDHQVIVNRIDYEM 178  
 241 IKGEDMSVLDVQEMLSLGVIPEDSEVIRSTRGPFVLYNKPPTLALAGLAPQAMRL 300  
 179 TKKGMILDSIDITLRLSVKLGAVPDDNITVSTNKGPIVLDEKAS--AGCAFNRIRRI 237

301 VEGDSMAVWVEEPKRGFS 322  
 238 IGED---VPIMDLNTEHQIGFS 256

## RESULT 11

097JM4 PRELIMINARY; PRT; 263 AA.  
 097JM4  
 DT 01-OCT-2001 (Tremblrel 18, Created)  
 DT 01-OCT-2001 (Tremblrel 18, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel 20, Last annotation update)  
 DE Septum site-determining protein MIND, ATPase.  
 GN CACT249.  
 OS Clostridium acetobutylicum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Clostridiales; Clostridiaceae; Clostridium.  
 OX NCBI\_TaxID=1488;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 824 / VKM B.1787,  
 RX MEDLINE=21359325; PubMed=1146286,  
 RA McElling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
 RA Gibson R., Lee H.M., Dubois J., Liu D., Hitti J., Wolf Y.I.,  
 RA Tatusov R.L., Sabache P., Doucette-Stamm L., Soucaille P., Daly M.J.,  
 RA Bennett G.N., Koonin E.V., Smith D.P.;  
 RT "Genome sequence and comparative analysis of the solvent-producing  
 RT bacterium Clostridium acetobutylicum".  
 RL J. Bacteriol. 183:4823-4838(2001).  
 DR EMBL, AE007638; AAK7922.1;  
 DR InterPro: IPR000707; ATPase\_Para  
 DR Pfam: PF00991; Para; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 263 AA; 28769 MW; B06AFBEA9DCFD3D73 CRC64;

Query Match 33.8%; Score 559.5; DB 16; Length 263;  
 Best Local Similarity 44.7%; Pred. No. 4.6e-38;  
 Matches 117; Conservative 57; Mismatches 77; Indels 11; Gaps 4;

61 VITSGKGVGKTTTANVGLSLARYGSVAIDADLGLRNLDLLGLERNVYTCVE 120  
 5 VITSGKGVGKTTTANVGLSLARYGSVAIDADLGLRNLDLLGLERNVYTCVE 64  
 121 NGDCRIDQALVPRKPSNFELLCISPRSKLPMGFGKALEMLVDALKTRPESSPDEI 180  
 65 ENRCRTKQALIRKRYENHLRLPASQTRDK-----DALPEVEKEVVGK 117  
 181 DCPAGIDAGFITAITPANEAVALTTPDITALRDADRVGLGECGIDRIKMIIVNPTVM 240  
 119 DCPAGIEGFEFNAIVGADRAIVNPEVTSVRDADRVIGKIDAKGIEDHQVIVNRIDYEM 178  
 241 IKGEDMSVLDVQEMLSLGVIPEDSEVIRSTRGPFVLYNKPPTLALAGLAPQAMRL 300  
 179 VKRGMILGIEDVILNLAIKLIGVVPNDKQITVSTNKGPIVLNQNAN--AGKAFRIARVR 237

## RESULT 12

092BH0 PRELIMINARY; PRT; 266 AA.  
 092BH0  
 DT 01-DEC-2001 (Tremblrel 19, Created)  
 DT 01-DEC-2001 (Tremblrel 19, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel 21, Last annotation update)  
 DE MIND protein.  
 GN MIND OR LIN1579.  
 OS Listeria innocua.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Clostridiales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1642;

```

RN [1]
RP SEQUENCE FROM N A
PC STRAIN=C11P 11262 / SEPOVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Charbit A., Cherouani F., Couve E., de Darvar A., Deboux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duseigneur O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf U., Jackson D.,
RA Jones L.-M., Kaerst H., Kretz J., Kuhn M., Kunst F., Karpik G.,
RA Madueno E., Maltournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueder T., Simoes N., Tietzer A.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001);
DR EMBL, AL591613, CAC96810.1, -;
DR Listlist: L1N01579; -;
DR InterPro: IPR000707; ATPase_Para.
DR Pfam: PF00991; Para; 1.
KW Complete proteome.
SQ SEQUENCE 266 AA; 29074 MW; D45807D1EC2B4E97 CRC64,
Query Match 33.0%; Score 546.5; DB 16; Length 266;
Best Local Similarity 43.3%; Pred. No. 5; Se-37;
Matches 117; Conservative 60; Mismatches 80; Indels 13; Gaps 6;
CY 61 VITSGKGVGKTTTANVSLAPYGFSSVAIDADLGLNLLGLNRYNYTCEVI 120
DB 5 IVITSGKGVGKTTSTANLGTALALGCKVCLIDMDIGLRDLVGLNRIIYDLVDV 64
CY 121 NGDCRLDQALVDPKMSNFELLCISRPSPKLPWFGKALEMLVNAKYTPREGSPDFII 180
DB 65 EGPCKTHQAMTPDKRPDL-LFLPAACTTGRVAVSGEQMELINQL-PPD-YDFILI 119
CY 181 DCPAGIDAGFTATTPANEAVLTPPDITALLPADPVTLGLEDGIPDIPMIVNVPRTDM 240
DB 120 DCPAGIEGTGYKNAVAGADKAIIVTTPEISAVRADRIICLLKEDIPEPKLIINRIOTM 179
CY 241 IKGEIMMSVLVDQEMLSLGLVPEDESEVIRSTNGFPLVINKPPTLAGLAEQAARKL 300
DB 180 MNGGVMDIDETHTLSIELGLIIDDDEVIPSNSGDPVAM-LPNNRASQGYPIINAPRI 238
CY 301 VEQDSMKAVMVEEPKRGFFS---FFQG 326
DB 239 LGE---SIPLMSEAKRGFFRPLKQLFGG 265

```

## RESULT 13

```

OBY6Y7
AC OBY6Y7; PRELIMINARY; PRT; 266 AA.
ID OBY6Y7;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mind protein.
GN MIND OR LMO1544.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N A.
PC STRAIN=EGD-E / SEPOVAR 1/2A;
RX MEDLINE=21537279, PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Chabib A., Chetouani F., Couve E., de Darvar A., Deboux P.,
RA Dmanin F., Dominguez-Bernal G., Duchaud E., Durant L., Duseigneur O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Kaerst H., Kretz J., Kuhn M., Kunst F., Karpik G.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Karpik G.,

```

```

RA Madueno E., Maltournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueder T., Simoes N., Tietzer A.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001);
DR EMBL, AL591613, CAC96810.1, -;
DR Listlist: LMO01544; -;
DR InterPro: IPR000707; ATPase_Para.
DR Pfam: PF00991; Para; 1.
KW Complete proteome.
SQ SEQUENCE 266 AA; 29066 MW; CCEB1DE7A7F278A CRC64,
Query Match 33.0%; Score 545.5; DB 16; Length 266;
Best Local Similarity 43.9%; Pred. No. 6; Se-37;
Matches 115; Conservative 59; Mismatches 79; Indels 9; Gaps 5;

```

```

CY 61 VITSGKGVGKTTTANVSLAPYGFSSVAIDADLGLNLLGLNRYNYTCEVI 120
DB 5 IVITSGKGVGKTTSTANLGTALALGCKVCLIDMDIGLRDLVGLNRIIYDLVDV 64
CY 121 NGDCRLDQALVDPKMSNFELLCISRPSPKLPWFGKALEMLVNAKYTPREGSPDFII 180
DB 65 EGPCKTHQAMTPDKRPDL-LFLPAACTTGRVAVSGEQMELINQL-PPD-YDFILI 119
CY 181 DCPAGIDAGFTATTPANEAVLTPPDITALLPADPVTLGLEDGIPDIPMIVNVPRTDM 240
DB 120 DCPAGIEGTGYKNAVAGADKAIIVTTPEISAVRADRIICLLKEDIPEPKLIINRIOTM 179
CY 241 IKGEIMMSVLVDQEMLSLGLVPEDESEVIRSTNGFPLVINKPPTLAGLAEQAARKL 300
DB 180 MNGGVMDIDETHTLSIELGLIIDDDEVIPSNSGDPVAM-LPNNRASQGYPIINAPRI 238
CY 301 VEQDSMKAVMVEEPKRGFFS 322
DB 239 LGE---SIPLMSEAKRGFFRPLKQLFGG 257

```

## RESULT 14

```

O9K8H9
AC O9K8H9; PRELIMINARY; PRT; 264 AA.
ID O9K8H9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Septum site-determining protein.
GN MIND OR BH3027.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacteriia; Firmicutes; Bacillaceae; Bacillus.
OX NCBI_TaxID=6665;
RN [1]
RP SEQUENCE FROM N A.
PC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582, PubMed=1105832,
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki P., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kunara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL, AP001517, BAB06746.1, -;
DR InterPro: IPR000707; ATPase_Para.
DR Pfam: PF00991; Para; 1.
KW Complete proteome.
SQ SEQUENCE 264 AA; 28818 MW; 78D523BE583172F CRC64,
Query Match 32.4%; Score 536.5; DB 16; Length 264;
Best Local Similarity 41.1%; Pred. No. 3; Se-36;
Matches 109; Conservative 66; Mismatches 83; Indels 7; Gaps 3;
CY 61 VITSGKGVGKTTTANVSLAPYGFSSVAIDADLGLNLLGLNRYNYTCEVI 120

```

Job time : 91 secs

```

Db      5 IVITSGKGVGKTTTANIGTALALSGKKVCLVDITDGLRNDLVVMGLRNRIIVDLDVV 64
QY      121 NGDCRLDQALVDFKPSNFEELICISPPSKLPMFGGALMLVDALKTRPEGSPDFIIT 180
Db      65 EGRCRILKQALLIKDKRECECLNLLPAQOTDK--SAVTPEQMEIVEELKQ----EYDVL 118
QY      181 DCPAGIDAGFTAITPANEAVLVTPDITALPDADRVYTGLECDGIRIDIKMIVNRVPTDM 240
Db      119 DCPAGIEGCFKAVAGADKAVIVTTPPEISSVADADRIITGLEKEVEAPRLVNRIRGHM 178
QY      241 IKGEDMMSVLDVQEMIGSLGVPIDPESEVIRSTNRGFPVLNRPPTLAGIAFEQAAARL 300
Db      179 MKNGEMLDVDEIVSLIATIELLIGIVDDEVNIKFSNKGEPIALH--PDSKASVAVNRNARI 237
QY      301 VEODSMKAVMVEEPKKGFFSFFG 325
Db      234 LGFTVPLMSFFQEKGVLAIRFSFFG 262

```

## RESULT 15

```

067033
ID      067033      PRELIMINARY;      FRT;      262 AA.
AC      067033;
DT      01-AUG-1998 (TRENBLrel. 07, Created)
DT      01-AUG-1998 (TRENBLrel. 07, last sequence update)
DE      01-JUN-2002 (TRENBLrel. 21, last annotation update)
DE      Septum site-determining protein MIND.
CN      MIND2 OR AQ_877.
OS      Aquifex aeolicus.
OC      Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX      NCBI_TaxID=63363;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=VF5.
FX      MEDLINE=98196666; PubMed=9537320;
RA      Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA      Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjaj M., Huber R.,
RA      Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT      "The complete genome of the hyperthermophilic bacterium Aquifex
RT      aeolicus.";
RL      Nature 392:353-358(1998).
EMBL; AE000712; AAC06996.1; -.
DR      InterPro; IPR000707; ATPase_Para.
DR      Pfam; PF00991; Para; 1.
KW      Complete proteome.
SQ      SEQUENCE 262 AA; 28913 MW; 535C2E5F3D5B352A CRC64;

```

Query Match 31.2%; Score 516; DB 16; Length 262;

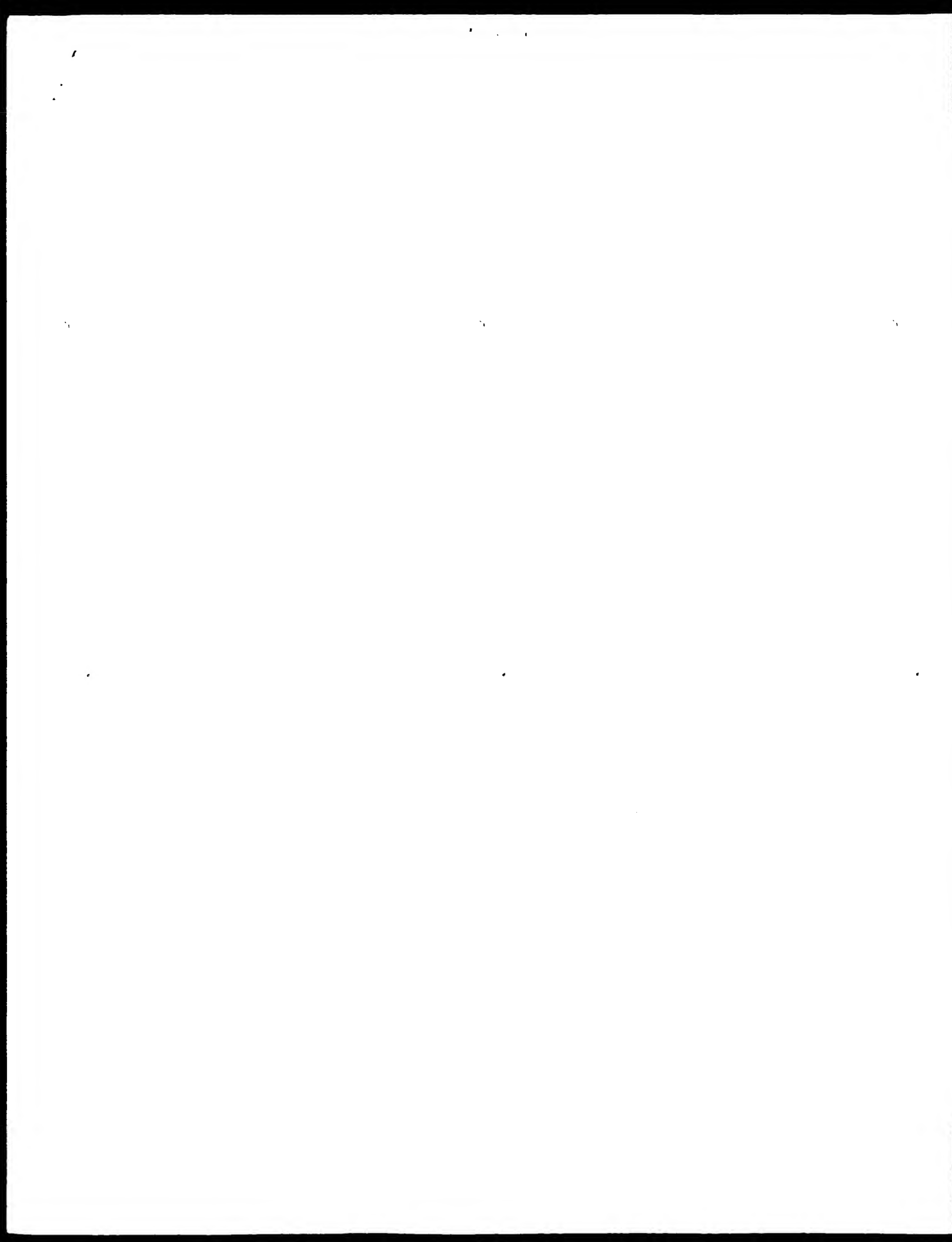
Best Local Similarity 41.4%; Pred. No. 1.7e-34;

```

Matches 108; Conservative 65; Mismatches 70; Indels 18; Gaps 3;
QY      60 IVITSGKGVGKTTTANIGTALALSGKKVCLVDITDGLRNDLVVMGLRNRIIVDLDVV 119
Db      4 VVITSGKGVGKTTTANIGTALALSGKKVCLVDITDGLRNDLVVMGLRNRIIVDLDVV 63
QY      120 INGDRLDQALVDFKPSNFEELICISPPSKLPMFGGALMLVDALKTRPEGSPDFIIT 179
Db      64 LEGRVPEKALVYKRGSLWLLPANQRANKVDIE---KMKVETIKNSGNYDYL 119
QY      180 IDCPAGIDAGFTAITPANEAVLVTPDITALPDADRVYTGLECDGIRIDIKMIVNRVPTDM 239
Db      120 VDSFAGIEGCFKAVAGADKAVIVTTPPEISSVADADRIITGLEKEVEAPRLVNRIRGHM 179
QY      240 IKGEDMMSVLDVQEMIGSLGVPIDPESEVIRSTNRGFPVLNRPPTLAGIAFEQAAARL 297
Db      180 MKRGAMLSVEDIVILAKAETIIGIPEEPKLVDFTRNGEPIVLDKFPASQAIT----- 233
QY      298 WRLVEODSMKAVMVEEPKKGFFSFFG 318
Db      234 -----DTARRLMGESIPDKR 248

```

Search completed: April 16, 2003, 09:07:46





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 09:05:27 ; Search time 20 seconds  
(Without alignments)  
1232.506 Million cell updates/sec

Title: US-10-067-989-1  
Perfect score: 1654  
Sequence: 1 MASRLRSTNHQSLLPSSL.....KAVWEDEPKKRGFSFEGG 326

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 288829 seqs, 7561385 residues

Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1654	100.0	326	12 US-10-067-989-1	Sequence 1, Appl1
2	792	47.9	284	12 US-10-067-989-2	Sequence 2, Appl1
3	643	38.9	266	12 US-10-067-989-3	Sequence 3, Appl1
4	479	29.0	270	12 US-10-067-989-4	Sequence 4, Appl1
5	197.5	11.9	279	9 US-09-738-626-6935	Sequence 6923, Ap
6	190	11.5	390	9 US-09-712-363-192	Sequence 192, App
7	176	10.6	288	9 US-10-102-806-730	Sequence 730, App
8	173.5	10.5	375	9 US-09-738-626-4744	Sequence 4744, Ap
9	164.5	9.9	478	9 US-09-738-626-3883	Sequence 3883, Ap
10	145.5	8.8	253	10 US-09-815-242-10915	Sequence 10915, A
11	134	8.1	264	10 US-09-815-242-11559	Sequence 11559, A
12	128.5	7.8	225	10 US-09-767-041-14	Sequence 14, Appl1
13	118	7.1	261	9 US-09-738-626-5070	Sequence 5070, Ap
14	107.5	6.5	308	10 US-09-791-171-94	Sequence 94, Appl1
15	107.5	6.5	666	10 US-09-791-171-70	Sequence 70, Appl1
16	102	6.2	261	10 US-09-841-132-517	Sequence 517, App
17	101.5	6.1	340	9 US-09-738-626-5233	Sequence 5233, Ap
18	99.5	6.0	262	10 US-09-815-242-5087	Sequence 5087, Ap
19	94	5.7	6304	9 US-10-147-026-16	Sequence 16, Appl1

20	91	5.5	644	9 US-09-948-137A-2	Sequence 2, Appl1
21	91	5.5	644	10 US-09-815-242-10318	Sequence 10318, A
22	90.5	5.5	335	9 US-09-738-626-6510	Sequence 6510, Ap
23	90.5	5.5	341	9 US-09-738-626-3835	Sequence 3835, Ap
24	89	5.4	453	10 US-09-815-242-10263	Sequence 10263, A
25	88	5.3	612	10 US-09-925-301-1358	Sequence 1358, Ap
26	87	5.2	421	10 US-09-815-242-11192	Sequence 11192, A
27	86.5	5.2	416	10 US-09-860-846-39	Sequence 39, Appl1
28	86.5	5.2	416	10 US-09-861-289-39	Sequence 39, Appl1
29	84	5.1	297	9 US-10-098-807-5	Sequence 5, Appl1
30	83.5	5.0	996	9 US-09-738-626-3847	Sequence 3847, Ap
31	83.5	5.0	3354	9 US-10-160-758-11	Sequence 11, Appl1
32	83.5	5.0	3354	9 US-10-160-758-12	Sequence 12, Appl1
33	83	5.0	229	10 US-09-767-041-13	Sequence 13, Appl1
34	83	5.0	357	9 US-09-712-363-280	Sequence 280, App
35	83	5.0	501	10 US-09-815-242-14089	Sequence 14089, A
36	83	5.0	577	9 US-09-738-626-4578	Sequence 4578, Ap
37	83	5.0	644	10 US-09-815-242-13760	Sequence 13760, A
38	83	5.0	1256	9 US-10-047-542-90	Sequence 90, Appl1
39	82.5	5.0	1080	10 US-09-781-558-4	Sequence 4, Appl1
40	82.5	5.0	1115	9 US-10-106-534-7	Sequence 7, Appl1
41	82.5	5.0	1135	10 US-09-737-149-35	Sequence 35, Appl1
42	82	5.0	258	10 US-09-954-314-12	Sequence 12, Appl1
43	82	5.0	369	9 US-09-712-363-170	Sequence 170, App
44	82	5.0	457	10 US-09-815-242-11948	Sequence 11948, A
45	82	5.0	473	9 US-09-738-626-3630	Sequence 3630, Ap

## ALIGNMENTS

```
RESULT 1
US-10-067-989-1
; Sequence 1, Application US/10067989
; Patent No. US20020144309A1
; GENERAL INFORMATION:
; APPLICANT: Dinkins, Randy
; APPLICANT: Reddy, M.S. Srinivasa
; APPLICANT: Collins, Glenn B.
; TITLE OF INVENTION: Transgenic plants expressing Mind or Mine and an efficient
; FILE OF INVENTION: method for plant chloroplast transformation and gene expression
; FILE REFERENCE: 028750-219
; CURRENT APPLICATION NUMBER: US/10/067, 989
; PRIOR APPLICATION NUMBER: US/02-05-24
; PRIORITY FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-067-989-1
Query Match 100.0%; Score 1654; DB 12; Length 326;
Best Local Similarity 100.0%; Pred. No. 7.8e-152;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASRLRSTNHQSLLPSSLQKTLISPRFVNNSRSPRSPIRSVQLQFRKDELAGEETRI 60
Db 1 MASRLRSTNHQSLLPSSLQKTLISPRFVNNSRSPRSPIRSVQLQFRKDELAGEETRI 60
QY 61 VITSGGKGGKTTTANVGLSLARGESVAIDADLGRNIDLLGLENNVNTCEVI 120
Db 61 VITSGGKGGKTTTANVGLSLARGESVAIDADLGRNIDLLGLENNVNTCEVI 120
QY 121 NGCDRLDQALVRDKRNSFELCISKPRSKLPMGCGKALEVVDALKTREGSPDFIT 180
Db 121 NGCDRLDQALVRDKRNSFELCISKPRSKLPMGCGKALEVVDALKTREGSPDFIT 180
QY 181 DCPAGIDAGITTAITPANEAVLTPDITALRDADRVTGLLECDGIRDKIIVNRVTD 240
Db 181 DCPAGIDAGITTAITPANEAVLTPDITALRDADRVTGLLECDGIRDKIIVNRVTD 240
```

```
OY 241 IKGEDMMSVLDVQEMIGSLGCVIPEDSEVIRSTNRGFPVLVINKPPTLAGIAFEQAAWRL 300
DB 241 IKGEDMMSVLDVQEMIGSLGCVIPEDSEVIRSTNRGFPVLVINKPPTLAGIAFEQAAWRL 300
OY 301 VEQDSMKAVMVEEPPKRGFFSFGG 326
DB 301 VEQDSMKAVMVEEPPKRGFFSFGG 326
```

RESULT 2  
US-10-067-989-2

```
; Sequence 2, Application US/10067989
; Patent No. US20020144309A1
; GENERAL INFORMATION:
; APPLICANT: Dinkins, Randy
; APPLICANT: Reddy, M.S. Srinivasa
; APPLICANT: Collins, Glenn B.
; TITLE OF INVENTION: Transgenic plants expressing mind or mine and an efficient
; FILE REFERENCE: 028750-219
; CURRENT APPLICATION NUMBER: US/10/067,989
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/267,488
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Chlorella vulgaris
US-10-067-989-2
```

Query Match 47.9%, Score 792, DB 12, Length 284;  
Best Local Similarity 60.4%, Pred. No. 1,3e-68;  
Matches 165; Conservative 39; Mismatches 55; Indels 14; Gaps 5;

```
OY 59 RIVVITSGKGVGKTTTANVGLSLARYGSVAIDADLGRNLDLLGLENRVNYTCVE 118
DB 19 RIVVITSGKGVGKTTTANVGLSLARYGSVAIDADLGRNLDLLGLENRVNYTCVE 118
OY 119 VINGCDRLDQALVDRKMSNFELLCISPKRSKLPWGFGKALEMVLDAKTRPGSPDFI 178
DB 79 IVEGQRLDQALVDRKMSNFELLCISPKRSKLPWGFGKALEMVLDAKTRPGSPDFI 178
OY 179 IIDCPAGIDAGFTAITPANEAVLVTPDITLADPDRVVTGLIEDCDGIPDIKMYNVRVT 238
DB 134 IIDCPAGIDAGFTAITPANEAVLVTPDITLADPDRVVTGLIEDCDGIPDIKMYNVRVT 238
OY 239 DMKEDMMSVLDVQEMIGSLGCVIPEDSEVIRSTNRGFPVLVINKPPTLAGIAFEQAAW 298
DB 194 DMKEDMMSVLDVQEMIGSLGCVIPEDSEVIRSTNRGFPVLVINKPPTLAGIAFEQAAW 298
OY 299 RLV-EQDSMKAVMVEEPPKRGFFSFGG 326
DB 254 RLV-EQDSMKAVMVEEPPKRGFFSFGG 326
```

RESULT 3  
US-10-067-989-3

```
; Sequence 3, Application US/10067989
; Patent No. US20020144309A1
; GENERAL INFORMATION:
; APPLICANT: Dinkins, Randy
; APPLICANT: Reddy, M.S. Srinivasa
; APPLICANT: Collins, Glenn B.
; TITLE OF INVENTION: Transgenic plants expressing mind or mine and an efficient
; FILE REFERENCE: 028750-219
; CURRENT APPLICATION NUMBER: US/10/067,989
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/267,488
; PRIOR FILING DATE: 2001-02-09
```

```
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Symnechocystis
US-10-067-989-3
```

Query Match 38.9%, Score 643; DB 12; Length 266;  
Best Local Similarity 52.8%, Pred. No. 3e-54;  
Matches 130; Conservative 51; Mismatches 59; Indels 6; Gaps 2;

```
OY 59 RIVVITSGKGVGKTTTANVGLSLARYGSVAIDADLGRNLDLLGLENRVNYTCVE 118
DB 3 RIVVITSGKGVGKTTTANVGLSLARYGSVAIDADLGRNLDLLGLENRVNYTCVE 118
OY 119 VINGCDRLDQALVDRKMSNFELLCISPKRSKLPWGFGKALEMVLDAKTRPGSPDFI 178
DB 63 VINGCDRLDQALVDRKMSNFELLCISPKRSKLPWGFGKALEMVLDAKTRPGSPDFI 178
OY 179 IIDCPAGIDAGFTAITPANEAVLVTPDITLADPDRVVTGLIEDCDGIPDIKMYNVRVT 238
DB 117 IIDCPAGIDAGFTAITPANEAVLVTPDITLADPDRVVTGLIEDCDGIPDIKMYNVRVT 238
OY 239 DMKEDMMSVLDVQEMIGSLGCVIPEDSEVIRSTNRGFPVLVINKPPTLAGIAFEQAAW 298
DB 177 DMKEDMMSVLDVQEMIGSLGCVIPEDSEVIRSTNRGFPVLVINKPPTLAGIAFEQAAW 298
OY 299 RLV-EQDSMKAVMVEEPPKRGFFSFGG 326
DB 237 RLV-EQDSMKAVMVEEPPKRGFFSFGG 326
```

RESULT 4  
US-10-067-989-4

```
; Sequence 4, Application US/10067989
; Patent No. US20020144309A1
; GENERAL INFORMATION:
; APPLICANT: Dinkins, Randy
; APPLICANT: Reddy, M.S. Srinivasa
; APPLICANT: Collins, Glenn B.
; TITLE OF INVENTION: Transgenic plants expressing mind or mine and an efficient
; FILE REFERENCE: 028750-219
; CURRENT APPLICATION NUMBER: US/10/067,989
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/267,488
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-067-989-4
```

Query Match 29.0%, Score 479; DB 12; Length 270;  
Best Local Similarity 39.9%, Pred. No. 2.1e-38;  
Matches 113; Conservative 58; Mismatches 82; Indels 30; Gaps 8;

```
OY 59 RIVVITSGKGVGKTTTANVGLSLARYGSVAIDADLGRNLDLLGLENRVNYTCVE 118
DB 3 RIVVITSGKGVGKTTTANVGLSLARYGSVAIDADLGRNLDLLGLENRVNYTCVE 118
OY 119 VINGCDRLDQALVDRKMSNFELLCISPKRSKLPWGFGKALEMVLDAKTRPGSPDFI 178
DB 63 VINGCDRLDQALVDRKMSNFELLCISPKRSKLPWGFGKALEMVLDAKTRPGSPDFI 178
OY 179 IIDCPAGIDAGFTAITPANEAVLVTPDITLADPDRVVTGLIEDCDGIPDIKMYNVRVT 238
DB 117 IIDCPAGIDAGFTAITPANEAVLVTPDITLADPDRVVTGLIEDCDGIPDIKMYNVRVT 238
OY 239 DMKEDMMSVLDVQEMIGSLGCVIPEDSEVIRSTNRGFPVLVINKPPTLAGIAFEQAAW 298
DB 231 DMKEDMMSVLDVQEMIGSLGCVIPEDSEVIRSTNRGFPVLVINKPPTLAGIAFEQAAW 298
```

Db 178 LLRTYNGPVSFGMLSMEDVLELRLKLVGVLPEDDSVLPASNGEPVILINDAD-AG 236  
 QY 291 LAFEQAMPLEVDSDMKAVMEED-----KKGFFS-PFGG 326  
 Db 237 KAV-----ADVEPLDGEPPFPFIEFEKKGFLLPFGG 270

## RESULT 5

US-09-738-626-6925  
 ; Sequence 6925, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAMA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAKO  
 ; APPLICANT: SENO, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738, 626  
 ; PRIOR FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 6925  
 ; LENGTH: 279  
 ; TYPE: PRT  
 ; ORGANISM: Corynebacterium glutamicum  
 ; US-09-738-626-6925

Query Match 11.9%; Score 197 5; DR 9; Length 279;  
 Best local Similarity 26.4%; Pred. No. 3.6e-11;  
 Matches 71; Conservative 46; Mismatches 109; Indels 43; Gaps 10;  
 QY 58 PRIVITSGKGVGKTTTAVNGSLARYGSVAIDADGLRLDLLLGLNRPV-AYTC 116  
 Db 7 PLLITANQKGVGKTTSTVNLASLAHGLKLVVDLD-PQGNASTALGVEHRSGLTSS 65  
 QY 117 VEVINGDRLDQALVDRKWSNFELLCTSKPRSKLPMGFGKALEW-----LVDAI 167  
 Db 66 YELLGECTADEAM--QSTANENLFCI-----PATIDLAGAELELVSYRREYRLADAL 118  
 QY 168 KTR--PESSPFIITIDCAGIDAGFTTATIPANEAVLVTPIDTALRDADRVTGLECDG 225  
 Db 119 GREFLDKHDFPMIIDCPPIGLLTINAMTAVNEVLIPICEYVALGEGVQL-----LNN 173  
 QY 226 IRDIMGIVNR-----VRTDMIKGRDMMSVLDDVQEMGLSLLG-VIPEDSEVI 271  
 Db 174 ITMLKQHLNRQHLHSAILLTMYDARTNLAE---QVATEVNDHGDVLDLKNKIPRSYKVS 229  
 QY 272 RSTNRGFLVLRKPPTLGLAFEQAAWRL 300  
 Db 230 EAPGYGQTVIEYDFGSGAMAYLDAKEL 258

## RESULT 6

US-09-712-363-192  
 ; Sequence 192, Application US/09712363  
 ; Patent No. US20020164588A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eisenberg, David  
 ; APPLICANT: Rotstein, Sergio H.

APPLICANT: Marcotte, Edward M.  
 ; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
 ; FILE REFERENCE: 07419-032001  
 ; CURRENT APPLICATION NUMBER: US/09/712,363  
 ; PRIOR FILING DATE: 2000-11-13  
 ; PRIOR APPLICATION NUMBER: PCT/US00/02246  
 ; PRIOR FILING DATE: 2000-01-28  
 ; PRIOR APPLICATION NUMBER: 60/179,531  
 ; PRIOR FILING DATE: 2000-02-01  
 ; PRIOR APPLICATION NUMBER: 60/117,844  
 ; PRIOR FILING DATE: 1999-01-29  
 ; PRIOR APPLICATION NUMBER: 60/118,206,  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: 60/126,593  
 ; PRIOR FILING DATE: 1999-03-26  
 ; PRIOR APPLICATION NUMBER: 60/134,093  
 ; PRIOR FILING DATE: 1999-05-14  
 ; PRIOR APPLICATION NUMBER: 60/134,092  
 ; PRIOR FILING DATE: 1999-05-14  
 ; PRIOR APPLICATION NUMBER: 60/165,124  
 ; PRIOR FILING DATE: 1999-11-12  
 ; PRIOR APPLICATION NUMBER: 60/165,086  
 ; PRIOR FILING DATE: 1999-11-12  
 ; NUMBER OF SEQ ID NOS: 292  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 192  
 ; LENGTH: 390  
 ; TYPE: PRT  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; US-09-712-363-192

Query Match 11.5%; Score 190; DB 9; Length 390;  
 Best local Similarity 24.0%; Pred. No. 3e-10;  
 Matches 68; Conservative 58; Mismatches 117; Indels 40; Gaps 9;  
 QY 32 VNPSRSPRSVLOFN-RKPELAGETP-----RIVVITSGKGVGKTTTAVNGSLARY 86  
 Db 95 VMSDEQRLRLKQVRSPTREPVIIPFAQPSLTVAVVAGSGKGVGKSTVTLAAAMAVR 154  
 QY 87 GFSVAIDADGLRLDLLLGLNRPVYTCVEVINGDCRLDQALVDRKWSNFELLCTSK 146  
 Db 155 GLSIGVDADHIGHSIRPMKGTDRPQV-----BSMLPPIAHQVKVISAQ 202  
 QY 147 -PRSKLPMGFGKAL-----EMLVDALKTRPGSPFIITIDCAGID--AGFTTATIPAN 198  
 Db 203 FTQATPVVMKGFMLRALQGFADLVY---WCDDLVLILLDPFGVDVAISVAGLIPNA 258  
 QY 199 EAVLVTPIDTALPDADPVTGILFCQGIPIDEMTVNRVPTDMIKGEDMSVLD-----VQ 253  
 Db 259 ELVVTTFQCLAAAVARASSIALQTRQRTVSVVENKSGTLTLPVITWLVFEGGAGPLVA 318  
 QY 254 EMLG-----LSLLGVIPEDESEVIRSTNRGFLVLRKPPTLGL 290  
 Db 319 ERLSRVAGADVPLLGQIPLDPALVAGDSGVPLVSSPDSAG 361

## RESULT 7

US-10-102-806-730  
 ; Sequence 730, Application US/10102806  
 ; Publication No. US20030054421A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: P0103P1C1  
 ; CURRENT APPLICATION NUMBER: US/10/102,806  
 ; PRIOR FILING DATE: 2002-03-22  
 ; PRIOR APPLICATION NUMBER: 09/925,298  
 ; PRIOR FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05981  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 846  
 SOFTWARE: Patent Ver. 2.0  
 SEQ ID NO: 730  
 LENGTH: 288  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-103-806-730

Query Match 10.6%; Score 176; DB 9; Length 288;  
 Best Local Similarity 22.8%; Pred No 4.5e-09;  
 Matches 68; Conservative 53; Mismatches 135; Indels 42; Gaps 8.

QY 23 PFVVRPFSPSPFISVLGFNRKP-ELAGEFERIVITSGKGVKTTTANVGLSLARYG 87  
 DB 2 PAVGTIELCPRAFGGGKAAEPGNLAG-VPHIILVSGKGVKSTSTELALALRHAG 60  
 QY 88 FSVVAIDADLGLRLDLLLGLGNFVNVTGVVINQDCFLDQALFVDRKSNFELLCSKP 147  
 DB 61 KVVGIIIVTLGCPSPFPMIGAQGRAVHGQ-----DRGMARV-FLDREGS 103  
 QY 148 RSKLPMGF-----GKALEWLVDAIKTRPBGSPFIIDCPACID--AGFIT 192  
 DB 104 ISLMSVGFLEPKPRDAVVPMPGPKYNALIKCFVSDVAMGELDYLVVDPPTGSDSHMATIE 163  
 QY 193 AITPANE--AVLVTPPDITALPDADRVTGILECDGIPDTFWVNPVPCTMIKGEKMSVL 256  
 DB 164 ALRPFQPLAIVVTPPAVSVGVDRBELTFPKYKGLRVMGIVEMSGFTCPHCTECTSVF 223  
 QY 251 -----DVOEMGLSLGLVPEDESEVIRSTNRGPFVLNKPPLLAGAFDQAMRLVE 302  
 DB 224 SRGGEELAQLAGVPLGSLVPLDPLMRTLEEGHDFIQEFGSPAPALSTIAQKIID 281

## RESULT 8

US-09-738-626-4744  
 Sequence 4744, Application US/09738626  
 Publication No. US20020197605A1  
 GENERAL INFORMATION:  
 APPLICANT: NAKAGAWA, SATOSHI  
 APPLICANT: MIZOGUCHI, HIROSHI  
 APPLICANT: ANDO, SEIKO  
 APPLICANT: HAYASHI, MIKIRO  
 APPLICANT: OCHIAI, KEIKO  
 APPLICANT: YOKOI, HARUHIKO  
 APPLICANT: TATEISHI, NAKO  
 APPLICANT: SENOH, AKIHIRO  
 APPLICANT: IKEDA, MASATO  
 APPLICANT: OZAKI, AKIO  
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 FILE REFERENCE: 249-125  
 CURRENT APPLICATION NUMBER: US/09/738, 626  
 CURRENT FILING DATE: 2000-12-18  
 PRIOR APPLICATION NUMBER: JP 99/377484  
 PRIOR FILING DATE: 1999-12-16  
 PRIOR APPLICATION NUMBER: JP 00/159162  
 PRIOR FILING DATE: 2000-04-07  
 PRIOR APPLICATION NUMBER: JP 00/280988  
 PRIOR FILING DATE: 2000-08-03  
 NUMBER OF SEQ ID NOS: 7059  
 SOFTWARE: Patent Ver. 3.0  
 SEQ ID NO: 4744  
 LENGTH: 375  
 TYPE: PRT  
 ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-4744

## Query Match

Best Local Similarity 10.5%; Score 173.5; DB 9; Length 375;  
 Matches 66; Conservative 53; Mismatches 88; Indels 65; Gaps 13.

QY 55 GETPRIVITSGKGVGKTTTANVGLSLARYGFSVVAIDADLGLRLDLLLGLGNVNY 114  
 DB 110 GSTTRVAVASGKGVGKSSMTVNLAALAKRGLSVGIIIDADIVGHSGVPMGLSDRPH- 168

QY 115 TCVEVINGDCFLDQALVDRKSNFELLCSK-PSKSLPMGFGKAL-----EWLVDAIK 168  
 DB 169 -----QVDD-MIMPPQAHGVKMSIAHFTEGNAPVVRGMLHRAIOQLTDVF- 216  
 QY 169 TRPBGSPFIIDCPACIDAGFITA--ITPANEAVLVTPPDITALRDAR 216  
 DB 217 ---WQDDILILLDPPTGDIATVAOLIPNAELLIVTTPQAAAAEVAERACTISVQTNQ 273  
 QY 217 -VTGILE-----CDG-IPDI-----PMIVNRVCTMIKGEKMSVLDQEMGLSL 261  
 DB 274 KVACVIENMSAMVLPDGTMDVDFGTGGGKIADRLTA--VTGEE-----VKVI 319  
 QY 282 GVPEDESEVIRSTNRGPFVLNKP--FTLAGL 291  
 DB 320 GSVPLDPLRLRGDVGNPDAISEPHSPTAAAI 351

## RESULT 9

US-09-738-626-3883  
 Sequence 3883, Application US/09738626  
 Publication No. US20020197605A1  
 GENERAL INFORMATION:  
 APPLICANT: NAKAGAWA, SATOSHI  
 APPLICANT: MIZOGUCHI, HIROSHI  
 APPLICANT: ANDO, SEIKO  
 APPLICANT: HAYASHI, MIKIRO  
 APPLICANT: OCHIAI, KEIKO  
 APPLICANT: YOKOI, HARUHIKO  
 APPLICANT: TATEISHI, NAKO  
 APPLICANT: SENOH, AKIHIRO  
 APPLICANT: IKEDA, MASATO  
 APPLICANT: OZAKI, AKIO  
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 FILE REFERENCE: 249-125  
 CURRENT APPLICATION NUMBER: US/09/738, 626  
 CURRENT FILING DATE: 2000-12-18  
 PRIOR APPLICATION NUMBER: JP 99/377484  
 PRIOR FILING DATE: 1999-12-16  
 PRIOR APPLICATION NUMBER: JP 00/159162  
 PRIOR FILING DATE: 2000-04-07  
 PRIOR APPLICATION NUMBER: JP 00/280988  
 PRIOR FILING DATE: 2000-08-03  
 NUMBER OF SEQ ID NOS: 7059  
 SOFTWARE: Patent Ver. 3.0  
 SEQ ID NO: 3883  
 LENGTH: 478  
 TYPE: PRT  
 ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-3883

## Query Match

Best Local Similarity 9.9%; Score 164.5; DB 9; Length 478;  
 Matches 58; Conservative 34; Mismatches 68; Indels 43; Gaps 10;

QY 41 IPSVQFNRKPELAGFPPIVITSGKGVKTTTANVGLSLARYGFSVVAIDADLRLP 100  
 DB 230 LRTNQLFN-----VGGSSVVFVSSANPGGKSTTSVNLALAEGRSVALLLEADRLRP 305  
 QY 101 NLDLLGLGNFVNVTGVVINQDCFLDQALVDRKSNFELLCSKSLPMG---FG 156  
 DB 306 FVSKYLGVEGNAGLT--DILIGKAEVNDVL--QPMWGRQLVYL--PAGPIPPNSELG 358  
 QY 157 GKALEWLVDAIKTRPBGSPFIIDCPAGI-----DAGFITAIT-----PANE 199  
 DB 359 SAEMEKVIAEL---EESFDVVIIDAPPALAVTDAVIGHGAGILIAVAGSTKKPELE 414  
 QY 200 AVLVTTPDITLRDAD-RVTGLL 221  
 DB 415 ATLSL-----LENADANVGVV 431

## RESULT 10

```
US-09-815-242-10915
; Sequence 10915, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10915
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10915

Query Match      8.8%; Score 145 5; DB 13; Length 253;
Best Local Similarity 20.1%; Pred. No. 3.3e-06;
Matches 55; Conservative 51; Mismatches 72; Indels 95; Gaps 12;

QY 59 RIVITSGKGVGKTTTANVGLSLARYGFSVAIDAD-----LGLNR-----LD 103
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 3 RIVSVNKGKGVGKTTTANVGLSLARYGFSVAIDAD-----LGLNR-----LD 62
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 104 LILGLE-----NPVNY-----TCVEVINGDCRLDQALVDPYWSNFELLCISPRSY 150
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 63 VLVNEEPITSVVQHTSPENLDIVPATITQAGAFETELTSMWRESF----- 107
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 151 LPMGGKALEVLDALTPPEGSPDFIILDCPAGIDAGFITATTPANEAVLVTPTDITA 210
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 108 -----LKAIDVVR-----DMYDFVLIDCPSPSLGHLTINAFI-ASDSILP----- 147
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 211 LRDADRVTLGLED--GIPDIKMIIVPVF-----TMIKGED--MMSVLVQEMGLSL 261
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 148 -----VQCEYYALEGLSOLINTIRLVQKHPNPELKBGLMVMYARNLGAEEV 197
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 262 -----GVIPDSEVIRSTNGFPLV 281
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 198 EEVRKYFFPKYVDTIIPNVRLSEAPSHGLPII 230
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
US-09-815-242-11559
; Sequence 11559, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
```

```
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11559
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11559

Query Match      8.1%; Score 134; DB 10; Length 264;
Best Local Similarity 23.3%; Pred. No. 4.6e-05;
Matches 63; Conservative 45; Mismatches 80; Indels 82; Gaps 13;

QY 60 IVVITSGKGVGKTTTANVGLSLARYGFSVAIDAD-----LGLNRDLGLLENR 111
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 5 IIVANQKGVGKTTTANVGLSLARYGFSVAIDAD-----LGLNRDLGLLENR 56
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 112 VNYTCVEVINGDCRLDQALV-----ENKWSNFELLCISK-----FSKLPWGF 156
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 57 IDYIVHWLIRKQISQVILKTPFLDVPVSLGAGFEKTFYDSQDENKKEGL----- 111
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 157 GKALEMVDALKTPRESF-----DFIILDCPAGIDAGFITATTPANEAVLVTPTDIT 212
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 112 -----KLNAL-----GSVVKLIDYITILISFPALGPLTINSLSAHSVIRP----- 152
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 213 DADPVTGLLED--GIPDIKMIIVPVF-----TMIKGEDMMSVLVQEMGLSLGV 264
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 153 -----IQCEFPALGCTLLMLTIRK-CKSTNPKLRGFLPTMHWVQ LNT-KQVL 202
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 265 PE-----DSEVIRSTNGFPLVANKPEPTLA 289
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 203 AELFKYFDSFEFFDSATGATGEYIMIPKSVKLA 232
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
US-09-767-041-14
; Sequence 14, Application US/09767041
; Patent No. US2002005168A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Hilida
; TITLE OF INVENTION: STREPTOCOCCUS SUE VACCINES AND DIAGNOSTIC TESTS
; FILE REFERENCE: 2183-4726
; CURRENT APPLICATION NUMBER: US/09/767,041
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: PCT/NL99/00460
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: EP98202465.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: EP98202467.1
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 53
```

SOFTWARE: PatentIn version 3.0  
 SEQ ID NO 14  
 LENGTH: 225  
 TYPE: PRT  
 ORGANISM: Streptococcus suis  
 FEATURE:  
 NAME/KEY: misc feature  
 OTHER INFORMATION: CPS2C  
 US-09-767-041-14

Query Match 7.8%; Score 128.5; DB 10; Length 225;  
 Best Local Similarity 23.2%; Pred. No. 0 conts;  
 Matches 56; Conservative 31; Mismatches 61; Indels 93; Gaps 9;

QY 52 ELAGETPRIVITSGKGVGKTTTANVGLSLARYGSVAIDDLGRN-----L 102  
 DB 29 QLSGADIKVGTIVSKVSEKSTTAASLAIAVARGVKTIVLADAD--IRNSVMPGFEP 86  
 QY 103 PLLGLIENPVNTVCVEVINGCPLD-----QALVPDKPMSNFELLGISK 146  
 DB 87 TITGLITVLAET-TDLSCGLCTDIPMLVYESGKVSFNTALLQSF--NFENLLATL 142  
 QY 147 PRSKLPMGFGKALEMLVDALKTREPGSPDFIIDCP---AGIDAGFTAITPANEAVLV 203  
 DB 143 RRYV-----DYVIDCPPLGLVIDAIIIAQKCDAMVAVV- 176  
 QY 204 TTPRITALRDAFVTVGLLETCGIPDIKMTVNPVPTDMKCGEDMMSVLDVQEMGLSLGV 263  
 DB 177 -----EAGNVKCSLSLKKVK-----EOLQOTCTPPLGV 203  
 QY 264 I 264  
 DB 204 I 204

## RESULT 13

US-09-738-626-5070  
 Sequence 5070, Application US/09738626  
 Publication No. US20020197605A1  
 GENERAL INFORMATION:

APPLICANT: NAKAGAMA, SATOSHI  
 APPLICANT: MIZOGUCHI, HIROSHI  
 APPLICANT: ANDO, SEIKO  
 APPLICANT: HAYASHI, MIKIRO  
 APPLICANT: OCHIAI, KEIKO  
 APPLICANT: YOKOI, HARUHIKO  
 APPLICANT: TATEISHI, NAOKO  
 APPLICANT: SENOH, AKIHIRO  
 APPLICANT: IKEDA, MASATO  
 APPLICANT: OZAKI, AKIO  
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 FILE REFERENCE: 249-125  
 CURRENT APPLICATION NUMBER: US/09/738,626  
 PRIOR FILING DATE: 2000-12-18  
 PRIOR APPLICATION NUMBER: JP 99/377484  
 PRIOR FILING DATE: 1999-12-16  
 PRIOR APPLICATION NUMBER: JP 00/159162  
 PRIOR FILING DATE: 2000-04-07  
 PRIOR APPLICATION NUMBER: JP 00/280988  
 PRIOR FILING DATE: 2000-08-03  
 NUMBER OF SEQ ID NOS: 7059  
 SOFTWARE: PatentIn ver 3.0  
 SEQ ID NO 5070  
 TYPE: PRT  
 LENGTH: 261  
 ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-5070

Query Match 7.1%; Score 118; DB 9; Length 261,  
 Best Local Similarity 27.8%; Pred. No. 0.0016;  
 Matches 54; Conservative 31; Mismatches 67; Indels 42; Gaps 9;

QY 60 IVVITSGKGVGKTTTANVGLSLARYGSVAIDAD-----LGLNLDLLGLENR 111

DB 9 IIMANOKGVGKTTSTINAGCLAENGRVLLVLDLPOGALTAGIGHYD----- 60  
 QY 112 VNYTCVE-VINGDCPLDQALVR-----DKPMSNFELLGISKPRSLPMGFGKALEMLV 164  
 DB 61 VDIIVYDLAMVNNSTIDQAIHHTGLPDLVDVSPANIDL---SAAEIOLVNVEGRE-QTLA 115  
 QY 165 DALKTREPGSP-DFIIDCPAGIDAGFTAITPANEAVLVTFEDTLARD ... 213  
 DB 116 PAL--RPVWKDYDFIILDCCPSLGLLTVNALACAHGVIIIMECEYFSLPGLALITDTEVK 173  
 QY 214 ADPVTVGLECCGI 226  
 DB 174 VADRLNFDLETLGI 187

## RESULT 14

US-09-791-171-94  
 Sequence 94, Application US/09791171  
 Patent No. US20020094336A1  
 GENERAL INFORMATION:

APPLICANT: ANDERSEN, Peter  
 APPLICANT: NIELSEN, Rikke  
 APPLICANT: OETTINGER, Thomas  
 APPLICANT: RASMUSSEN, Peter Birk  
 APPLICANT: ROSENKRANDS, Ida  
 APPLICANT: WELDLING, Karin  
 TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS  
 FILE REFERENCE: 670001-2002.1  
 CURRENT APPLICATION NUMBER: US/09/791,171  
 PRIOR FILING DATE: 2001-02-20  
 PRIOR APPLICATION NUMBER: 09/050,739  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 0376/97  
 PRIOR FILING DATE: 1997-04-02  
 PRIOR APPLICATION NUMBER: 1277/97  
 PRIOR FILING DATE: 1997-11-10  
 PRIOR APPLICATION NUMBER: 60/044,624  
 PRIOR FILING DATE: 1997-04-18  
 PRIOR APPLICATION NUMBER: 60/070,488  
 PRIOR FILING DATE: 1998-01-05  
 NUMBER OF SEQ ID NOS: 173  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 94  
 LENGTH: 308  
 TYPE: PRT  
 ORGANISM: Mycobacterium tuberculosis  
 US-09-791-171-94

Query Match 6.5%; Score 107.5; DB 10; Length 308;  
 Best Local Similarity 21.6%; Pred. No. 0.021;  
 Matches 61; Conservative 46; Mismatches 97; Indels 79; Gaps 14;

QY 61 VVITSGKGVGKTTTANVGLSLARY-GFSVAIDDLGRN-----DLTL 106  
 DB 61 IAVVGLKKGAGKTLTALGSLAQVADRIILADDPAGNIALAVGQSGATLADVLA 120  
 QY 107 GLE-----NRVN-VTCVEVIN-----GDCPLDQALVDPKMSNFELLGISKPRSLPMGF 155  
 DB 121 EKELSHYNDIRAHTSVNAVNLVLEVLPAPEYSSAQRALSDADWH-----FIADPASRF----- 171  
 QY 156 GSKALEMLVDALKTREPGSPDFIIDCPAG-----IDAGFTAITPANEAVLVTFEDTLARD 211  
 DB 172 -----YNLVLADCGAGFDPPLTRGVSTV--SGVVVAASVSDIGA 209  
 QY 212 PDANPVTGLECDGIPDIK---MIVNPVPTDMKEDMMSVLDV GEMLGLSLGV 263  
 DB 210 QQASVALDMLPNMYGLASRACVNIHI---MPSEPNVAVEDLVRHHECCGVQGRVVV 265  
 QY 264 IPDESEVIRSTNRGF---PLVLNKPPTLAGLA---FEQDAMR 299



